

ALEXANDRE MONTE



**CONCORDÂNCIA VERBAL E VARIAÇÃO: UM ESTUDO
DESCRITIVO-COMPARATIVO DO PORTUGUÊS
BRASILEIRO E DO PORTUGUÊS EUROPEU**



ALEXANDRE MONTE

**CONCORDÂNCIA VERBAL E VARIAÇÃO: UM ESTUDO
DESCRITIVO-COMPARATIVO DO PORTUGUÊS
BRASILEIRO E DO PORTUGUÊS EUROPEU**

Tese apresentada ao Programa de Pós-Graduação em Linguística e Língua Portuguesa da Faculdade de Ciências e Letras da Universidade Estadual Paulista “Júlio de Mesquita Filho”, *Campus* de Araraquara, como requisito para obtenção do título de Doutor em Linguística e Língua Portuguesa.

Orientadora: Profa. Dra. Rosane de Andrade Berlinck

Araraquara
2012

Monte, Alexandre

Concordância verbal e variação: um estudo descritivo-comparativo do português brasileiro e do português europeu. – 2012

171 f. ; 30 cm

Tese (Doutorado em Linguística e Língua Portuguesa) – Universidade Estadual Paulista, Faculdade de Ciências e Letras, *Campus* de Araraquara

Orientadora: Profa. Dra. Rosane de Andrade Berlinck

1. Linguística. 2. Língua portuguesa. 3. Língua portuguesa – Concordância. I. Título.

ALEXANDRE MONTE

**CONCORDÂNCIA VERBAL E VARIAÇÃO: UM ESTUDO
DESCRITIVO-COMPARATIVO DO PORTUGUÊS
BRASILEIRO E DO PORTUGUÊS EUROPEU**

BANCA EXAMINADORA

Orientadora: Profa. Dra. Rosane de Andrade Berlinck
UNESP – Araraquara

Profa. Dra. Silvia Rodrigues Vieira
UFRJ – Rio de Janeiro

Profa. Dra. Juliana Bertucci Barbosa
UFTM – Uberaba

Prof. Dr. Sebastião Carlos Leite Gonçalves
UNESP – São José do Rio Preto

Profa. Dra. Beatriz Nunes de Oliveira Longo
UNESP – Araraquara

Araraquara, 05 de novembro de 2012

Aos meus pais, Ana Maria e Antonio, pelo amor incondicional e pela dedicação. A vocês, devo esta conquista.

Aos meus sobrinhos, Juninho, Giovanna e Lara.

AGRADECIMENTOS

Em primeiro lugar, a Deus, pelo verdadeiro Amor.

À minha grande família, em especial à minha mãe, Ana Maria, e ao meu pai, Antonio, pela torcida constante, pelo carinho e pelas orações. A família é a primeira escola do amor.

À Profa. Dra. Rosane de Andrade Berlinck, por ter aceitado ser, novamente, a minha orientadora. Obrigado pela confiança, atenção, dedicação, amizade e pela orientação competente e dialógica. O nosso convívio, desde a graduação, tem sido um fator determinante para o meu crescimento profissional e pessoal.

À Profa. Dra. Maria Filomena Candeias Gonçalves, supervisora do meu trabalho no exterior, pela acolhida e por todo apoio para realizar a pesquisa em Évora/Portugal.

À Profa. Dra. Silvia Rodrigues Vieira, à Profa. Dra. Juliana Bertucci Barbosa, ao Prof. Dr. Sebastião Carlos Leite Gonçalves e à Profa. Dra. Beatriz Nunes de Oliveira Longo, pela leitura atenta do trabalho e pelas valiosas sugestões.

Aos meus irmãos (Fabio e Rodrigo) e às minhas cunhadas (Vanessa e Antonise), pelo carinho e convívio amigável.

Aos meus amigos fiéis (Nalva, Osmair, Pedro (Pedrinho), Rosana, Rosemary (Rosinha), Sabrina Amorim (Sá), Vanessa (Piu), Zelma), pela generosidade, pelo amor e pelo ombro amigo nos momentos difíceis. Em especial, ao Osmair, à Rosinha e à Sá. “Um amigo fiel é um forte refúgio. Quem o achou, achou um tesouro. Um amigo fiel não tem preço. É incalculável o seu valor. Um amigo fiel é um bálsamo para a vida” (Eclo 6, 14-16a).

Aos amigos que conheci na escola Deriggi e na escola Degan, aos amigos de São Carlos e de Araraquara, pelo companheirismo, pela amizade e por todos os momentos de descontração.

Aos amigos, colegas e professores do Programa de Pós-Graduação e a todos os integrantes do NEVAR da UNESP de Araraquara – *Núcleo de Estudos sobre Variação Linguística* – coordenado pela Profa. Dra. Rosane de Andrade Berlinck, pelos ensinamentos e pelos diálogos enriquecedores.

Aos amigos que fiz em terras lusitanas, pelos momentos “giros”.

À Profa. Dra. Angela Cecília de Souza Rodrigues e à Profa. Dra. Cássia Regina Coutinho Sossolote, pelo incentivo.

À Profa. Dra. Ângela Balça, à Socióloga Dores Correia, ao Pe. Manuel da Silva Ferreira, ao Prof. Dr. Paulo Costa, ao Prof. Paulo Tapadas e à Profa. Regina Camacho, por toda atenção e pelo apoio em Évora/Portugal.

À Profa. Dra. Ernestina Carrilho, docente da Universidade de Lisboa, pelos diálogos sobre a variação da concordância verbal no português europeu.

Aos funcionários da Seção de Pós-Graduação e da Biblioteca da UNESP de Araraquara, pelo trabalho competente.

Aos meus informantes são-carlenses e eborenses, pela disponibilidade e generosidade. Sem eles, não conseguiria concretizar este trabalho.

À CAPES, pelo financiamento do estágio de Doutorado no exterior, no período de fevereiro a dezembro de 2010, junto à Universidade de Évora, em Portugal.

Valeu a pena? Tudo vale a pena
Se a alma não é pequena.
Quem quer passar além do Bojador
Tem que passar além da dor.
Deus ao mar o perigo e o abismo deu,
Mas nele é que espelhou o céu.

Fernando Pessoa

MONTE, Alexandre. **Concordância verbal e variação**: um estudo descritivo-comparativo do português brasileiro e do português europeu. 2012. Tese (Doutorado em Linguística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, São Paulo, Brasil.

RESUMO

O objetivo geral deste trabalho é descrever a variação na concordância verbal de terceira pessoa do plural no português brasileiro (PB) e no português europeu (PE) e discutir o estatuto dessa variação nas duas variedades. É um estudo norteado pelos pressupostos teórico-metodológicos da Teoria da Variação e Mudança Linguística (WEINREICH, LABOV & HERZOG, 1968; LABOV, 1972, 1994, 2001, 2003). Por meio de um estudo empírico e sincrônico, analisamos a variação na concordância verbal de terceira pessoa do plural na fala de 18 pessoas residentes na cidade de São Carlos, localizada no interior do Estado de São Paulo/Brasil, e na fala de 18 pessoas residentes na cidade de Évora, situada no Alentejo, sul de Portugal. Essas pessoas foram estratificadas em função do sexo e nível de escolaridade. Os nossos resultados revelam uma significativa diferença quantitativa entre as duas variedades. No PB, com 48,2% (686/1.422) de concordância verbal, o fenômeno se configura como uma regra variável. Já no PE, com 93,1% (1.340/1.440) de concordância verbal, a regra parece ter *status* semicategórico. Para o tratamento estatístico dos dados, utilizamos o programa Goldvarb-X, que realiza uma análise multivariada. Na variedade brasileira, todos os grupos de fatores linguísticos e sociais controlados foram atuantes: escolaridade, saliência fônica verbal, tipo estrutural do sujeito/SN, paralelismo formal no nível oracional, gênero/sexo, posição do sujeito/SN em relação ao verbo, traço semântico do sujeito/SN, tipo de verbo (verbo 'ser' *versus* outros verbos), distância entre o sujeito/SN e o verbo em número de sílabas e presença ou ausência do 'que' relativo ou complementizador. Na variedade europeia, foram apenas seis os grupos de fatores selecionados: posição do sujeito/SN em relação ao verbo, traço semântico do sujeito/SN, tipo estrutural do sujeito/SN, tipo de verbo (verbo 'ser' *versus* outros verbos), saliência fônica verbal e gênero/sexo. As diferenças quantitativas e qualitativas encontradas mostram que há uma regra mais consistente do ponto de vista estrutural no PE para a concordância verbal de terceira pessoa do plural.

Palavras-chave: Concordância verbal. Português brasileiro. Português europeu. Variação linguística. Língua falada.

MONTE, Alexandre. **Verb agreement and variation**: a descriptive-comparative study of Brazilian Portuguese and European Portuguese. 2012. Thesis (Ph.D. in Linguistics and the Portuguese Language) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, São Paulo, Brazil.

ABSTRACT

The main objective of this thesis is to describe the variation in third-person plural verb agreement in Brazilian Portuguese (BP) and European Portuguese (EP) and discuss the statute of this variation in both varieties. The study is guided by the theoretical and methodological assumptions of the Theory of Language Variation and Change (WEINREICH, LABOV & HERZOG, 1968; LABOV, 1972, 1994, 2001, 2003). Through an empirical and synchronic study, we analyzed the variation in third-person plural verb agreement in the speech of 18 people who live in São Carlos, a city in the state of São Paulo/Brazil, and in the speech of 18 inhabitants of Évora, a city located in Alentejo, southern Portugal. These people were stratified by gender and educational level. Our results show a significant quantitative difference between the two varieties. In BP, with 48.2% (686/1.422) of verb agreement, the phenomenon is a variable rule. On the other hand, in EP, with 93.1% (1.340/1.440) of verb agreement, the rule seems to have a semi-categorical status. For the statistical treatment of the data, we used the Goldvarb-X program, which performs a multivariate analysis. In the Brazilian variety, all groups of linguistic and social factors included were active: schooling, verbal phonic salience, subject/NP structural type, formal parallelism in clause level, gender/sex, position of the subject/NP in relation to the verb, subject/NP semantic feature, type of verb ('to be' and others), distance between subject/NP and verb in number of syllables and presence or absence of the relative or complementizer 'que' (that/who/which). In the European variety, only six groups of factors were selected: position of the subject/NP in relation to the verb, subject/NP semantic feature, subject/NP structural type, type of verb ('to be' and others), verbal phonic salience and gender/sex. The quantitative and qualitative differences that were found show that there is a more consistent rule from a structural point of view in EP for the third-person plural verb agreement.

Keywords: Verb agreement. Brazilian Portuguese. European Portuguese. Linguistic variation. Spoken language.

SUMÁRIO

INTRODUÇÃO	11
1 FUNDAMENTAÇÃO TEÓRICA	16
1.1 Teoria da variação e mudança.....	16
1.2 Problemas empíricos para a teoria da mudança linguística.....	22
1.3 Sobre as origens do português brasileiro.....	24
2 CONCORDÂNCIA VERBAL NO PB E NO PE: UMA REVISÃO	30
2.1 No Português do Brasil.....	30
2.2 No Português Europeu.....	37
3 O UNIVERSO DESTE ESTUDO	42
3.1 As comunidades de fala	42
3.1.1 Considerações sobre a cidade de São Carlos.....	42
3.1.2 Considerações sobre a cidade de Évora.....	50
3.2 Procedimentos metodológicos da pesquisa	56
3.2.1 Os <i>corpora</i> sob análise.....	56
3.2.2 A coleta dos dados.....	59
3.2.3 A seleção dos dados.....	62
3.2.4 Critérios de exclusão dos dados.....	66
3.2.5 Grupos de fatores (ou variáveis).....	70
3.2.5.1 Grupos de fatores linguísticos (ou variáveis linguísticas).....	71
3.2.5.2 Grupos de fatores sociais (ou variáveis sociais).....	85
4 ANÁLISE DOS DADOS: APRESENTANDO E DISCUTINDO OS RESULTADOS	88
4.1 Notas introdutórias	88
4.2 Análise e discussão dos resultados gerais do PE	90
4.2.1 Posição do sujeito/SN em relação ao verbo.....	91
4.2.2 Traço semântico do sujeito/SN.....	92
4.2.3 Tipo estrutural do sujeito/SN.....	97
4.2.4 Tipo de verbo: verbo 'ser' <i>versus</i> outros verbos.....	99
4.2.5 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	103
4.2.6 Gênero.....	106
4.2.7 Distância entre o sujeito/SN e o verbo em número de sílabas.....	107
4.2.8 Presença ou ausência do 'que' relativo ou complementizador.....	107
4.2.9 Paralelismo formal no nível oracional (marcas no sujeito).....	108
4.2.10 Escolaridade.....	110
4.3 Análise e discussão dos resultados do PE sem as ocorrências com o verbo 'ser'	112
4.3.1 Posição do sujeito/SN em relação ao verbo.....	113
4.3.2 Traço semântico do sujeito/SN.....	113
4.3.3 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	114
4.3.4 Escolaridade.....	115

4.4 Análise e discussão dos resultados gerais do PB	117
4.4.1 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	118
4.4.2 Tipo estrutural do sujeito/SN.....	121
4.4.3 Paralelismo formal no nível oracional (marcas no sujeito).....	123
4.4.4 Posição do sujeito/SN em relação ao verbo.....	126
4.4.5 Traço semântico do sujeito/SN.....	127
4.4.6 Tipo de verbo: verbo 'ser' <i>versus</i> outros verbos.....	128
4.4.7 Distância entre o sujeito/SN e o verbo em número de sílabas.....	131
4.4.8 Presença ou ausência do 'que' relativo ou complementizador.....	132
4.4.9 Gênero.....	134
4.4.10 Escolaridade.....	135
4.5 Análise e discussão dos resultados do PB sem as ocorrências com o verbo 'ser'	141
4.5.1 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	142
4.5.2 Posição do sujeito/SN em relação ao verbo.....	143
4.5.3 Traço semântico do sujeito/SN.....	143
4.5.4 Escolaridade.....	144
4.6 Comparação dos resultados do PB e do PE	146
4.6.1 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural.....	146
4.6.2 Tipo estrutural do sujeito/SN.....	148
4.6.3 Posição do sujeito/SN em relação ao verbo.....	150
4.6.4 Traço semântico do sujeito/SN.....	151
4.6.5 Tipo de verbo: verbo 'ser' <i>versus</i> outros verbos.....	152
4.6.6 Gênero.....	153
4.6.7 Escolaridade.....	154
5 CONSIDERAÇÕES FINAIS	157
6 REFERÊNCIAS	163

INTRODUÇÃO

O objetivo desta investigação é descrever a variação na concordância verbal de terceira pessoa do plural no português falado da cidade de São Carlos, localizada no interior do estado de São Paulo/Brasil, e no português falado da cidade de Évora, situada no Alentejo, sul de Portugal, e discutir o estatuto dessa variação nas duas variedades.

No Brasil, a concordância verbal é um fenômeno variável que atrai muito a atenção social e, de acordo com Bagno (2009, p. 104), é o fenômeno linguístico que mais desperta atitudes de discriminação entre os falantes urbanos letrados. Por outro lado, em uma entrevista concedida à revista *Discutindo Língua Portuguesa*, a linguista portuguesa Maria Helena Mira Mateus (2006) comenta que, em Portugal, a distância entre o que se poderia chamar o português popular (ainda que não se use essa denominação) e as variedades de prestígio não é tão grande como no Brasil. Por isso, esta pesquisa visa compreender as diferenças e semelhanças entre o português brasileiro (doravante PB) e o português europeu (doravante PE) no âmbito da concordância verbal de 3ª pessoa do plural, correlacionando a variação existente em cada comunidade de fala estudada com fatores estruturais da língua e com fatores sociais (gênero e escolaridade).

Ao focalizar a relação entre a variação e a escolaridade, pretendemos avaliar o papel desse fator em duas realidades diferentes e, desse modo, compreender como se estabelece tal relação. A especificidade da conjuntura sócio-histórico-econômica que caracteriza cada uma dessas realidades provavelmente esteja associada a diferentes padrões de valoração linguística. Nosso estudo, desse modo, pode permitir uma melhor compreensão quanto à atuação de parâmetros sociais nos processos de variação e mudança.

De acordo com Lemle e Naro (1977, p. 40), ao abordar a variação, é possível captar alguns fatores que determinam a mudança linguística, além de revelar os caminhos pelos quais uma mudança se difunde e possibilitar um melhor entendimento da própria história do desenvolvimento das línguas.

Além da contribuição teórico-descritiva, não podemos deixar de ressaltar a relevância social deste trabalho, pois a concordância verbal no Brasil, como já mencionamos, é um dos fenômenos variáveis da língua que mais tem sofrido

avaliação social e alimentado manifestações de preconceito e intolerância linguísticas.

Em todos os estudos analisados no Brasil, pudemos constatar que é possível correlacionar a aplicação variável de concordância entre sujeito/SN e verbo tanto a fatores internos (linguísticos) quanto a fatores externos (sociais). Estamos seguros de que a concordância verbal é um fenômeno linguístico que não pode ser analisado apenas em termos de suas relações internas na gramática, mas deve ser visto como parte de um contexto sociocultural mais amplo. Dessa forma, nosso estudo foi norteado pelos pressupostos teórico-metodológicos da “Teoria da Variação e Mudança Linguística” (WEINREICH, LABOV e HERZOG, 1968; LABOV, 1972, 1994, 2001, 2003).

Esse fenômeno começou a ser estudado no Brasil na década de setenta por Miriam Lemle e Anthony Naro e, desde então, inúmeros trabalhos já foram realizados em diversas comunidades de fala de nosso país (cf. LEMLE e NARO, 1977; NARO e LEMLE, 1977; NARO, 1981; RODRIGUES, 1987; NARO e SCHERRE, 1991, 1999a, 1999b, 2000, 2003a, 2003b, 2010; SCHERRE e NARO, 1993, 1998a, 1998b, 2000; SCHERRE, NARO e CARDOSO, 2007; VIEIRA, 1995; MONGUILHOTT, 2001; PEREIRA, 2004; GAMEIRO, 2005, 2009; MONTE, 2007; RUBIO, 2008; LUCCHESI, BAXTER e SILVA, 2009; OLIVEIRA, 2010; entre outros). Esses estudos, como já salientamos, têm evidenciado que a variação da concordância verbo-sujeito/SN na terceira pessoa gramatical plural é controlada por diversas variáveis, entre as quais destacamos a saliência fônica da oposição singular/plural dos verbos; o traço semântico do sujeito/SN; a posição e a distância do sujeito/SN em relação ao verbo; a presença ou ausência do ‘que’ relativo ou complementizador; o paralelismo formal no nível oracional e discursivo; o tipo estrutural do sujeito/SN; o tipo de verbo; a faixa etária; a escolaridade; o gênero¹.

Já com dados do português europeu não há um grande número de estudos variacionistas que se ocupam da análise da concordância verbal de terceira pessoa do plural. Encontramos os trabalhos de Carrilho (2003), Varejão (2006), Naro e Scherre (2007), Mota e Vieira (2008), Monguilhott (2009), Gandra (2009), Almeida

¹ Explicaremos melhor as variáveis nos procedimentos metodológicos do nosso trabalho.

(2010), Bazenga (2011), Rubio (2012), Vieira (2012) e Brandão e Vieira (2012)². Os resultados das pesquisas realizadas com dados do PE revelam uma significativa diferença quantitativa em relação aos resultados das pesquisas realizadas com dados do PB. Segundo Varejão (2006), as questões postas aos linguistas interessados nesse campo podem ser resumidas à busca por resposta a uma questão de partida: qual é de fato a dimensão quantitativa e qualitativa dessas semelhanças e diferenças?

Scherre e Naro (2005, p. 63) afirmam que “o português brasileiro e o português europeu partilham estruturas semelhantes e que as estruturas variáveis de uso comum no Brasil também podem ser encontradas em Portugal, com menos intensidade, tanto hoje quanto antes da colonização do Brasil”.

Lucchesi (2006, p. 108), por outro lado, levando em consideração a pouca variação na concordância verbal no português europeu, levanta a seguinte questão: “esse nível irrisório de variação não poderia ser melhor analisado como acidentes de performance, não se configurando, portanto, um processo efetivo de variação estruturada?”.

Diante da relevância do tema para a descrição do português em termos sociolinguísticos, especialmente para a interpretação das variedades europeia e brasileira, estabelecemos a seguinte questão de pesquisa: *Em que medida o estatuto variável da concordância verbal de terceira pessoa do plural pode ser postulado para o PB e para o PE?* Os objetivos específicos relacionados ao objetivo geral são:

(i) Realizar um estudo descritivo-comparativo da variação da concordância verbal de 3ª pessoa do plural entre a comunidade de fala são-carlense (PB) e a eborense (PE);

(ii) Analisar as variáveis linguísticas que possam estar condicionando / determinando a variação no âmbito da concordância verbal de 3ª pessoa do plural no PB e no PE;

(iii) Analisar a atuação das variáveis sociais gênero e escolaridade no condicionamento da concordância verbal de 3ª pessoa do plural no PB e no PE;

² Os trabalhos de Carrilho (2003), Varejão (2006) e Gandra (2009) se valeram de dados do mesmo *corpus* (*Corpus Dialectal para o Estudo da Sintaxe* (CORDIAL-SIN) do Centro de Linguística da Universidade de Lisboa).

(iv) Comparar, na medida do possível, os resultados obtidos com os de outros estudos.

Estamos propondo com nossa investigação uma comparação efetiva, com base em *corpora* especialmente constituídos para tal, segundo rigorosos critérios metodológicos. Pretendemos, assim, trazer uma contribuição para a compreensão de um fenômeno que está no centro das discussões sobre os processos que levaram à constituição do português brasileiro: deriva, crioulização ou transmissão linguística irregular (cf. BAXTER, 1992; LUCCHESI, 2003, 2006, 2008; LUCCHESI e BAXTER, 2009; SCHERRE e NARO, 2005; NARO e SCHERRE, 2007).

O trabalho apresenta-se dividido em quatro seções. Na primeira, expomos a fundamentação teórica. Destacamos os conceitos da teoria laboviana, os cinco problemas empíricos apresentados por Weinreich, Labov e Herzog (2006 [1968]) e as ideias de vários pesquisadores sobre as origens do português brasileiro.

Na segunda seção, apresentamos uma revisão bibliográfica de alguns estudos realizados sobre a concordância verbal de terceira pessoa do plural no PB e no PE.

Na terceira seção, apresentamos o universo de nossa pesquisa: as comunidades de fala onde o fenômeno foi estudado e quais foram os procedimentos metodológicos adotados para que a investigação fosse desenvolvida. É nessa seção que esclarecemos como foi feita a seleção dos dados e apresentamos as variáveis linguísticas e sociais.

Por fim, na última seção, expomos a análise e a discussão dos resultados do português brasileiro e do português europeu. Na sequência, apresentamos as considerações finais a que se chega com esta investigação.

Cabe, ainda, ressaltar que se trata do primeiro estudo sociolinguístico laboviano realizado com dados de fala da cidade de Évora, da Região Alentejana de Portugal. A cidade de São Carlos já possui um estudo sociolinguístico realizado por nós (cf. MONTE, 2007), mas os dados aqui apresentados foram obtidos de uma nova amostra de língua falada, para uma comparação efetiva com os dados de Évora. Sendo assim, a presente pesquisa vem complementar e aprofundar o nosso trabalho anterior.

De acordo com Naro e Scherre (2007, p. 116), “estudos do uso da língua em seu contexto social em terras europeias se fazem necessários para que possam ser

feitas comparações verdadeiramente comparáveis”. Além disso, sabemos que não se pode negar a importância de descrever a variação linguística do português do Brasil e do português de Portugal no maior número e diversidade de comunidades.

1 FUNDAMENTAÇÃO TEÓRICA

A existência de variação e de estruturas heterogêneas nas comunidades de fala investigadas está certamente bem fundamentada nos fatos. É a existência de qualquer outro tipo de comunidade de fala que deve ser posta em dúvida. (LABOV, 2008 [1972], p. 238)³.

1.1 Teoria da variação e mudança

Nosso estudo foi realizado com base nos princípios teóricos da “Teoria da Variação e Mudança Linguística” (WEINREICH, LABOV e HERZOG, 1968; LABOV, 1972, 1994, 2001, 2003). Para esse modelo, a natureza variável da língua é um pressuposto fundamental, que orienta e sustenta a observação, a descrição e a interpretação do comportamento linguístico. A concepção de língua como um sistema heterogêneo constitui o ponto crucial da ruptura epistemológica que a Teoria da Variação e Mudança Linguística opera em relação ao modelo estruturalista, como vem expresso em seu texto fundador:

Muito antes de se poder esboçar teorias preditivas da mudança linguística, será necessário aprender a ver a língua – seja de um ponto de vista diacrônico ou sincrônico – como um objeto constituído de heterogeneidade ordenada.

Os fatos da heterogeneidade, até agora, não se harmonizaram bem com a abordagem estrutural da língua. (WEINREICH, LABOV e HERZOG, 2006 [1968], p.35).

Segundo Paiva e Duarte (2006), o ponto de ruptura estabelecido por Weinreich, Labov e Herzog, em relação aos modelos dialetológicos anteriores e aos modelos estruturalistas vigentes na época, está na concepção de língua como um sistema heterogêneo ordenado, condição *sine qua non* para o estudo da mudança linguística. A análise da variação é pertinente, na medida em que esta é interpretada como uma condição indispensável para entender a mudança linguística.

Weinreich, Labov e Herzog, no final do texto *Empirical Foundations for a Theory of Language Change*, de 1968, explicitam algumas coordenadas teóricas

³ The existence of variation and heterogeneous structures in the speech communities investigated is certainly well-established in fact. It is the existence of any other type of speech community that may be placed in doubt. (LABOV, 1972, p. 203).

sobre a natureza da mudança linguística que podem ser tomadas como centrais para sua proposta:

1. A mudança linguística não deve ser identificada com deriva aleatória procedente da variação inerente na fala. A mudança linguística começa quando a generalização de uma alternância particular num dado subgrupo da comunidade de fala toma uma direção e assume o caráter de uma diferenciação ordenada.
2. A associação entre estrutura e homogeneidade é uma ilusão. A estrutura linguística inclui a diferenciação ordenada dos falantes e dos estilos através de regras que governam a variação na comunidade de fala; o domínio do falante nativo sobre a língua inclui o controle destas estruturas heterogêneas.
3. Nem toda variabilidade e heterogeneidade na estrutura linguística implica mudança; mas toda mudança implica variabilidade e heterogeneidade.
4. A generalização da mudança linguística através da estrutura linguística não é uniforme nem instantânea; ela envolve a covariação de mudanças associadas durante substanciais períodos de tempo, e está refletida na difusão de isoglossas por áreas do espaço geográfico.
5. As gramáticas em que ocorre a mudança linguística são gramáticas da comunidade de fala. Como as estruturas variáveis contidas na língua são determinadas por funções sociais, os idioletos não oferecem a base para gramáticas autônomas ou internamente consistentes.
6. A mudança linguística é transmitida dentro da comunidade como um todo; não está confinada a etapas discretas dentro da família. Quaisquer descontinuidades encontradas na mudança linguística são os produtos de descontinuidades específicas dentro da comunidade, mais do que os produtos inevitáveis do lapso geracional entre pais e filhos.
7. Fatores linguísticos e sociais estão intimamente inter-relacionados no desenvolvimento da mudança linguística. Explicações confinadas a um ou outro aspecto, não importa quão bem construídas, falharão em explicar o rico volume de regularidades que pode ser observado nos estudos empíricos do comportamento linguístico. (WEINREICH, LABOV e HERZOG, 2006 [1968], p. 125-126).

Para Lucchesi (2004), os princípios empíricos propostos no texto permitem-nos resolver a oposição paradoxal entre estrutura e mudança. A mudança linguística não é vista como exterior ao sistema, mas parte integrante do seu caráter normalmente heterogêneo.

Como se pode observar, a variação é sistemática, não aleatória e constitui uma característica intrínseca da língua e fonte da mudança. Os estudos realizados por Labov têm como principal característica a análise de discursos concretos, objetivando descrever a gramática efetiva de uma determinada comunidade de fala, bem como depreender as relações entre padrões linguísticos e sociais. Para Labov

(1972), o termo “sociolinguística” é redundante, uma vez que não se pode conceber uma linguística que não seja social:

Este tipo de pesquisa tem sido às vezes rotulado de “sociolinguística”, embora este seja um uso um tanto enganoso de um termo estranhamente redundante. A língua é uma forma de comportamento social [...] (LABOV, 2008 [1972], p. 215)⁴.

A Sociolinguística atua nas fronteiras entre língua e sociedade, focalizando os empregos concretos da língua. Os fenômenos de variação linguística são condicionados, não só por fatores internos à estrutura linguística, mas também por fatores extralinguísticos, de natureza social, ligados ao próprio falante e à situação em que a comunicação se processa.

Em qualquer comunidade de fala, independentemente de seu tamanho, há uma variação considerável entre os indivíduos: as mulheres não falam como os homens, os avós falam de modo diferente dos filhos e dos netos, e assim por diante. Além disso, mesmo os indivíduos considerados em sua singularidade não estão limitados a uma única variedade da língua. Sabemos que a rede social de um indivíduo, constituída pelas pessoas com quem esse indivíduo interage nos diversos domínios sociais, também é um fator determinante das características de seu repertório sociolinguístico. Assim, os falantes adquirem as variedades linguísticas próprias à sua região, à sua classe social, etc. Por isso, a sociolinguística variacionista vem dizer que a variação linguística deve ser observada e estudada na comunidade de fala.

A clássica definição de Labov (1972) é a de que a comunidade de fala não é definida por nenhum acordo/contrato marcado no uso de elementos linguísticos, mas pela participação num conjunto de normas compartilhadas. Para Labov (1972), essas normas podem ser observadas em tipos claros de comportamento avaliativo e pela uniformidade de padrões abstratos de variação que são invariantes em relação aos níveis particulares de uso. Dizer que os membros de uma comunidade de fala compartilham um conjunto comum de normas linguísticas não quer dizer que todos falam exatamente da mesma maneira.

⁴ This type of research has sometimes been labelled as “sociolinguistics”, although it is a somewhat misleading use of an oddly redundant term. Language is a form of social behavior [...] (LABOV, 1972, p. 183).

Guy (2000), fundamentado na proposta de Labov, apresenta um resumo esclarecedor com três características importantes de comunidade de fala:

- características linguísticas compartilhadas; isto é, palavras, sons ou construções gramaticais que são usados na comunidade, mas não o são fora dela.
- densidade de comunicação interna relativamente alta; isto é, as pessoas normalmente falam com mais frequência com outras que estão dentro do grupo do que com aquelas que estão fora dele.
- normas compartilhadas; isto é, atitudes em comum sobre o uso da língua, normas em comum sobre a direção da variação estilística, avaliações sociais em comum sobre variáveis linguísticas. (GUY, 2000, p. 18).

De uma perspectiva geral, podemos descrever as variedades linguísticas a partir de dois parâmetros extralinguísticos básicos: a variação geográfica (ou diatópica) e a variação social (ou diastrática). A variação geográfica ou diatópica está relacionada às diferenças linguísticas distribuídas no espaço físico, observáveis entre falantes de origens geográficas distintas. A variação social ou diastrática, por sua vez, relaciona-se a um conjunto de fatores que têm a ver com a identidade dos falantes e também com a organização sociocultural da comunidade linguística. Além da variação diatópica e da variação diastrática, também encontramos na literatura linguística a variação diafásica (variação estilística) e a variação diamésica (associada ao uso de diferentes meios e veículos). Segundo Ilari e Basso (2006, p.181), “a variação diamésica compreende, antes de mais nada, as profundas diferenças que se observam entre a língua falada e a língua escrita”.

Das possíveis variáveis externas (extralinguísticas) pertinentes ao estudo da variação, as que mais têm sido discutidas são: o estilo de fala, o gênero, a idade, a escolaridade, a profissão, a classe social, a região ou zona de residência e a procedência do falante.

Incorporando a variação na descrição e na teoria linguísticas, Labov (1972, 1994, 2001, 2003) introduz alguns conceitos teórico-metodológicos de extrema importância para a pesquisa sociolinguística. Segundo o autor, todo sistema linguístico é dotado de um conjunto de regras que não podem ser violadas, sob pena de dificultar ou mesmo inviabilizar a compreensão dos enunciados. A esse conjunto de leis internas se costuma dar o nome de ‘regras categóricas’ (i.e. regras

linguísticas que sempre se aplicam). Mas, além das ‘regras categóricas’, existem as ‘regras semicategóricas’ e as ‘regras variáveis’.

No texto “Some Sociolinguistic Principles”, Labov (2003, p. 241-243) explica os três tipos de regras linguísticas e a frequência com que cada tipo de regra opera.

Tipo de regra	Frequência com que opera
I – Categórica	100%
II – Semicategórica	95-99%
III – Variável	5-95%

Segundo Labov (2003, p. 242), “é comum encontrar regras semicategóricas no início ou no final de uma mudança linguística em progresso, em que a forma é muito rara para ser notada quando quer que ocorra”⁵.

O conceito de ‘regra variável’ é utilizado para substituir a noção de regra opcional do Estruturalismo, na medida em que não pressupõe variação livre, mas, sim, sistemática. A respeito da variação livre, Fischer e Bright compartilham as mesmas concepções de Labov:

“Variação livre” é naturalmente uma denominação e não uma explanação, pois não nos mostra a origem das variantes e nem porque os falantes as usam em proporções divergentes. A variação livre é, antes, um meio de se excluir tais questões da esfera da pesquisa imediata. (FISCHER, 1974, p. 88).

[...] os sociolinguistas rompem incisivamente com uma tendência linguística: a de tratar as línguas como sendo completamente uniformes, homogêneas ou monolíticas em sua estrutura; sob este ponto de vista, que vem sendo reconhecido atualmente como pernicioso, as diferenças encontradas nos hábitos de fala de uma comunidade eram encobertas como “variação livre”. Uma das maiores tarefas da sociolinguística é demonstrar que na verdade tal variação ou diversidade não é “livre”, mas correlacionada a diferenças sociais sistemáticas. (BRIGHT, 1974, p. 18).

As formas linguísticas em variação em uma determinada comunidade de fala são denominadas ‘variantes linguísticas’. Estas são definidas como formas alternativas de se dizer a mesma coisa, em um mesmo contexto. Embora sejam idênticas em seu valor referencial, as variantes podem opor-se quanto ao seu

⁵ It is common to find Type II rules at the beginning or at the end of a linguistic change in progress, where the form is rare enough to be noticed whenever it occurs. (LABOV, 2003, p. 242).

significado social e/ou estilístico. Ao conjunto de variantes dá-se o nome de ‘variável linguística’.

No início da Sociolinguística Variacionista, os estudos labovianos se concentraram no nível fonológico. Lavandera (1984 [1978]) não vai considerar adequada a extensão do conceito de variável linguística, com variantes idênticas em seu valor referencial, para o nível morfológico e sintático.

Lo que pondré en duda en este trabajo es que se pueda abandonar la base de variación fonológica con clara equivalencia semántica, y llevar a cabo el mismo tipo de estudio de variación con unidades morfológicas y sintácticas para las que es necesario haber demostrado que significan “lo mismo” antes de tratarlas como datos de variación. (LAVANDERA, 1984 [1978], p. 41-42).

He señalado algunas de las dificultades metodológicas que percibo en la definición de las variables no fonológicas, y espero haber indicado cómo estas dificultades surgen del hecho de que la variación no fonológica afecta formas con significado. (LAVANDERA, 1984, [1978], p. 45).

Labov (1978), em sua resposta a Lavandera, continuou defendendo a sua posição teórica. Hoje, passados mais de trinta anos, vemos que não há motivos para deixar a variação morfossintática fora dos estudos sociolinguísticos de cunho laboviano. De acordo com Mattos e Silva (2008), há várias pesquisas da variação morfossintática no português do Brasil (os clíticos e sua mobilidade, a concordância verbal e nominal, uso de relativas padrão/não padrão, preenchimento do sujeito, entre outras) que excluem a dificuldade levantada por Lavandera, porque nesses temas o significado lógico e referencial das variantes é o mesmo, como prevê a teoria escolhida. A concordância verbal no português do Brasil, por exemplo, constitui precisamente uma ‘regra variável’, ou uma ‘variável linguística’, que abrange duas ‘variantes’: a presença ou a ausência de marca de plural no verbo.

Como podemos perceber, a sociolinguística não aceita a visão da variabilidade como um fato aleatório e insiste na necessidade de um controle sistemático e empírico dos fatores estruturais (internos) e sociais que motivam o uso de uma ou outra variante.

Os estudos sociolinguísticos têm mostrado que a variação não é de maneira alguma aleatória e caótica. Ao contrário, é altamente estruturada. A abordagem quantitativa revolucionou o estudo da língua, demonstrando que o comportamento

linguístico é ainda mais fortemente estruturado do que se havia suspeitado anteriormente. De acordo com Trask (2006):

É claro que os primeiros linguistas perceberam essa variação, mas eles se inclinaram a desqualificá-la, por entender que se tratava de um fato marginal e sem consequências, ou mesmo como um estorvo atravessado no caminho das boas descrições. Hoje, ao contrário, reconhecemos que a variação é uma parte integrante e essencial da língua, e que a ausência de variação é quase patológica. (TRASK, 2006, p. 278).

1.2 Problemas empíricos para a teoria da mudança linguística

São cinco os problemas empíricos apresentados por Weinreich, Labov e Herzog (2006 [1968]) para uma teoria da mudança linguística: o problema dos fatores condicionantes (*the constraints problem*), o problema da transição (*the transition problem*), o problema do encaixamento (*the embedding problem*), o problema da avaliação (*the evaluation problem*) e o problema da implementação (*the actuation problem*).

Ao tratar do problema dos condicionamentos, uma importante pergunta que deve nortear o trabalho do pesquisador é: que fatores determinam as mudanças possíveis e em que direção elas acontecem? Segundo WLH, é fundamental, para uma teoria da mudança, determinar o conjunto de mudanças possíveis e condições possíveis para a mudança. Que restrições determinam mudanças possíveis?

Dessa forma, estamos controlando, neste estudo, vários grupos de fatores linguísticos e dois sociais (gênero e escolaridade), a fim de verificar quais estão condicionando ou restringindo a possível mudança na concordância verbal de 3ª pessoa do plural na comunidade de fala são-carlense (PB) e na eborense (PE).

Ao buscar “a trilha pela qual uma mudança linguística está caminhando para se completar” (2006 [1968] p. 90), o pesquisador estará envolvido no problema da transição. De que maneira uma língua muda? Como a língua passa de um estágio a outro? Como identificar a trajetória entre duas etapas de uma mudança linguística?

Sobre o problema da transição, WLH dizem:

A mudança se dá (1) à medida que um falante aprende uma forma alternativa, (2) durante o tempo em que as duas formas existem em contato dentro de sua competência, e (3) quando uma das formas se torna obsoleta. A transferência parece ocorrer entre grupos de pares de faixas etárias levemente diferentes; todas as evidências empíricas reunidas até agora indicam que as crianças não preservam as características dialetais de seus pais, mas sim as do grupo de pares que domina seus anos pré-adolescentes. (WEINREICH, LABOV e HERZOG, 2006 [1968], p. 122).

O problema da transição também leva em conta a transmissão de regras de uma geração a outra. É possível realizar um estudo da mudança mediante a observação do comportamento linguístico de falantes em diversas faixas etárias. É a perspectiva que se convencionou denominar de 'tempo aparente'.

No caso de você prever um caso de variação que já projete uma mudança dentro do sistema, o fator faixa etária é de extrema importância. Na impossibilidade de fazer um estudo longitudinal (um acompanhamento dos falantes desde a adolescência até a idade madura) sobre a variável, a amostragem da comunidade em grupos etários diferentes lhe dará a dimensão procurada. (TARALLO, 2002, p. 47).

Segundo Paiva e Duarte (2006), Weinreich, Labov e Herzog (1968) rompem com as fronteiras entre sincronia e diacronia. O entrelace dos dois eixos permite, então, um passo teórico importante: as evidências da variação sincrônica passam a constituir um excelente laboratório para a compreensão de mudanças já completadas, ocorridas no passado.

O problema do encaixamento, segundo Labov (2008 [1972]), tem dois aspectos: a mudança é vista como encaixada numa matriz de outras mudanças linguísticas, e também como encaixada num complexo social, correlacionada com mudanças sociais. Sendo assim, temos o encaixamento na estrutura linguística e o encaixamento na estrutura social.

O estudo do problema do encaixamento vai lidar com as mudanças que se encontram associadas a outras mudanças de uma maneira não acidental/casual. O trabalho de Galves (1993), por exemplo, vai relacionar o enfraquecimento da concordância verbal com outras mudanças sintáticas ocorridas no Brasil no século XIX. Para a sua argumentação, Galves (1993) utiliza os resultados das pesquisas de Berlinck (1989), sobre a fixação da ordem SVO, e de Cyrino (1990), Duarte (1991) e

Nunes (1990,1991), sobre a modificação do sistema pronominal. Segundo a pesquisadora, o enfraquecimento da concordância verbal está relacionado às outras mudanças observadas: aumento do preenchimento do sujeito e enrijecimento da ordem SV. Essa ideia também está presente no trabalho de Monguilhott (2009).

O problema da avaliação coloca a necessidade de se saber de que maneira as pessoas de uma comunidade de fala avaliam determinada variante de um fenômeno variável. Por meio da avaliação, podemos perceber como os membros da comunidade de fala reagem à mudança. Para WLH (2006 [1968], p. 124), “o nível de consciência social é uma propriedade importante da mudança linguística que tem de ser determinada diretamente”.

Em relação à concordância verbal, sabemos que é um fenômeno que não está imune à estigmatização e o preconceito linguístico com as pessoas que não a realizam é muito forte na sociedade brasileira. De acordo com Roncarati (2008, p. 47-48), “não é possível discutir preconceito e prestígio linguísticos e temas a eles afetos, como testes de atitudes e crenças linguísticas, sem se levar em conta um dos pilares sustentadores da inquirição sociolinguística sobre a variação e a mudança linguística: o problema da avaliação linguística [...]”.

Por que uma determinada mudança ocorre num tempo e lugar particulares e não em outros? Com essa pergunta, estamos diante do problema da implementação. Segundo WLH (2006) [1968],

O processo global da mudança linguística pode envolver estímulos e restrições tanto da sociedade quanto da estrutura da língua. A dificuldade do enigma da implementação é evidente no número de fatores que influenciam a mudança [...] (WEINREICH, LABOV e HERZOG, 2006 [1968], p. 124).

1.3 Sobre as origens do português brasileiro

Como as línguas são organismos que se desenvolvem e transformam, esse fato foi-se dando no Brasil e em Portugal ao mesmo tempo, isto é, tanto aqui como lá a língua se foi desenvolvendo, ou alterando, como quiserem.

Desde que a corrente se tinha bifurcado, cada um dos veios novos começou a modificar-se à parte, independente um do outro. A língua não é hoje em Portugal a mesmíssima de 1500; não o é também no Brasil. (COELHO, 1967 [1881], p.119-120).

Serafim da Silva Neto (1976, p. 67) delinea três fases da história da língua portuguesa no Brasil, que visam ao estudo das várias composições da população brasileira. A primeira fase tem início em 1532, com a colonização, e vai até 1654, com a expulsão dos holandeses. A segunda fase, de 1654 a 1808, propiciou o aumento da emigração do Reino. Segundo o autor, “o elemento indígena, incompatível com a civilização europeia, vai rareando e desaparecendo, ao mesmo passo que cresce a influência dos brancos e dos negros”. A terceira fase começa em 1808, com a chegada da família real.

Sabemos, principalmente pelos estudos de Aryon Dall’Igna Rodrigues, que inúmeras línguas foram faladas no território brasileiro pelos povos indígenas. Mas, segundo Bagno (2011, p. 233), os pesquisadores ainda não conseguiram chegar a um consenso a respeito de alguma influência das línguas indígenas sobre o português brasileiro que não seja a contribuição ao léxico.

Adolfo Coelho (1967 [1881]) não acredita na influência gramatical das línguas indígenas no PB. Ao analisar dois materiais que tentam provar tal influência, diz Coelho:

Como se vê os dois escritores brasileiros, que têm um longo conhecimento do tupi-guarani, nenhum fato apresentam que prove a influência gramatical que admitem: as suas indicações resumem-se a fatos lexicológicos, alguns dos quais são muito contestáveis. (COELHO, 1967 [1881], p. 102).

Câmara Jr. (2004 [1972]) também afirma que nenhum fonema indígena e nenhuma categoria ou flexão verbal tupi deixou vestígio no português do Brasil. Mas, em relação às línguas africanas, Câmara Jr. (2004 [1972]) vai dizer que a situação foi diferente. O autor até assume um falar crioulo no PB, que explicaria certas inovações e simplificações, mas não nega a deriva. Talvez a ideia de Câmara Jr. (2004 [1972]) seja a de um semicrioulo ou leve crioulição, como se observa por suas afirmações:

Os escravos negros adaptaram-se ao português sob a forma de uma falar crioulo. Nos latifúndios, ou fazendas, da época colonial e do Império o contato dos senhores brancos com seus escravos negros foi intenso e estreito. As crianças eram confiadas aos cuidados de amas escravas, as chamadas “mães-pretas”, e devem ter tomado de início, sem sentir, elementos do português crioulo que elas usavam. Por aí se poderia explicar certas inovações e simplificações do português do Brasil em face do europeu. É claro, entretanto, que não se dariam mudanças fonológicas e gramáticas profundas sem correspondência com as próprias tendências estruturais da língua portuguesa. (CÂMARA JR., 2004 [1972], p. 117-118).

É Adolfo Coelho (1967 [1881], p. 43) quem primeiro lança a ideia de uma possível criouliização do PB, ao afirmar que “diversas particularidades características dos dialetos crioulos repetem-se no Brasil”.

A partir daí, muito já se falou a respeito do intenso contato do português brasileiro com as línguas africanas. O número de escravos trazidos da África foi altíssimo. Os estudiosos divergem quanto ao número, mas Mattos e Silva (2004) afirma que os africanos e afrodescendentes estavam no patamar de 60% da população do Brasil entre os séculos XVII e XIX.

Mesmo com esse massivo contato, a ideia de uma origem crioula no PB foi combatida por alguns estudiosos. Antenor Nascentes e Serafim da Silva Neto, por exemplo, tentam minimizar as diferenças entre o português do Brasil e o de Portugal. Entretanto, quando vamos às suas obras, encontramos:

Outro tanto já não se pode dizer da língua popular pois muitas deturpações que ela apresenta encontram-se no dialeto crioulo cabo-verdiano, no da Guiné, no das ilhas de S. Tomé, Príncipe e Ano-Bom, no de Angola e no da costa oriental da África e a eles podem ser atribuídas. (NASCENTES, 1960, p. 263).

No português brasileiro não há, positivamente, influência de línguas africanas ou ameríndias. O que há é cicatrizes da tosca aprendizagem que da língua portuguesa, por causa de sua mísera condição social, fizeram os negros e os índios. (SILVA NETO, 1976, p. 96-97).

Como podemos notar, a ideia de Silva Neto (1976) se aproxima da ideia de Baxter (1992) e Lucchesi (2003, 2006, 2008), com o conceito de ‘transmissão linguística irregular’.

Lucchesi (2006) afirma que existem muitos registros históricos de processos potenciais de criouliização da língua portuguesa no Brasil. Ele acredita que não se desenvolveu no Brasil, de forma representativa, uma variedade criouliizada da língua portuguesa, como aconteceu, por exemplo, no arquipélago de Cabo Verde e nas Ilhas de São Tomé e Príncipe. Mas defende que o contato entre línguas, principalmente com as africanas, desencadeou importantes mudanças na estrutura das variedades da língua portuguesa que viriam a se desenvolver nos segmentos da base da sociedade brasileira, por meio do que denomina processo de ‘transmissão linguística irregular’.

O conceito de ‘transmissão linguística irregular’ é apresentado por Lucchesi (2003) da seguinte forma:

Assim, a transmissão linguística irregular constitui um contínuo de níveis diferenciados de socialização/nativização de uma língua segunda, adquirida massivamente, de forma mais ou menos imperfeita, em contextos sócio-históricos específicos. A criouliização típica se situa no extremo desse contínuo, numa situação em que o acesso à língua alvo foi extremamente reduzido, desencadeando um processo de reestruturação linguística independente, cujo resultado é a formação de uma nova língua qualitativamente distinta da(s) língua(s) que forneceu(ram) os modelos primários para a aquisição/nativização. Porém, mesmo no setor crioulo do contínuo, a natureza dinâmica dos fatores externos que condicionam a transmissão linguística irregular define resultados diversos, alguns mais radicais, outros menos radicais. (LUCCHESI, 2003, p. 274).

Rejeitando a interpretação crioulista da origem do português brasileiro, Naro e Scherre (2007) defendem a hipótese de que o impulso motor do desenvolvimento do PB veio já embutido na deriva secular da língua de Portugal. “Se as sementes trazidas de lá germinaram mais rápido e cresceram mais fortes, é que as condições, aqui, mostraram-se mais propícias devido a uma confluência de motivos” (NARO e SCHERRE, 2007, p.48).

Os pesquisadores, garimpendo as origens do PB, mostram que os contatos linguísticos ocorridos no Brasil não chegaram a modificar a tipologia estrutural da língua. Consultaram vários trabalhos da dialetologia portuguesa e localizaram no português europeu as origens de uma série de traços do português brasileiro, que são considerados por alguns pesquisadores vestígios de criouliização ou semicriouliização.

O modelo que Naro e Scherre (2007, p. 69 e 85) assumem é o da confluência de motivações, sem criouliização prévia do português. Sendo assim, rejeitam também a proposta de ‘transmissão linguística irregular’ defendida por Baxter (1992) e Lucchesi (2003, 2006, 2008)⁶.

⁶ Em nota, Bortoni-Ricardo (2011, p. 28) diz que “Naro e Scherre (2007) posicionam-se contra a hipótese da criouliização e aduzem o conceito da transmissão linguística irregular”. Na verdade, como já vimos, o conceito de ‘transmissão linguística irregular’ é de Baxter (1992) e Lucchesi (2003, 2006, 2008), e não de Naro e Scherre.

Nossa conclusão é que o português moderno do Brasil é o resultado natural da deriva secular inerente na língua trazida de Portugal, indubitavelmente exagerada no Brasil pela exuberância do contato de adultos, falantes de línguas das mais diversas origens, e pela nativização desta língua pelas comunidades formadas por estes falantes. Nossa proposta assim se resume: CONFLUÊNCIA DE MOTIVAÇÕES, com umas mais fortes no início do processo histórico e outras mais relevantes em estágios históricos seguintes. (NARO e SCHERRE, 2007, p. 69, destaque no original).

Na mesma corrente de pensamento, encontramos Noll (2008). Para o autor, o ambiente existente no Brasil contraria os pressupostos necessários para a formação de línguas crioulas, devido à situação continental, relação populacional e difusão das línguas africanas. Além disso, Noll (2008) consulta várias declarações da literatura de viagem do século XIX e conclui que essas declarações depõem contra uma criouliização.

Por outro lado, Galves (2008) e Petter (2009) advogam, assim como Baxter e Lucchesi, em favor da influência das línguas africanas no PB.

Petter (2009) focaliza formas de português faladas na África, na área não-crioula (Angola e Moçambique), chamando atenção para o fato de que as diferentes situações de contato, em épocas diversas, mas envolvendo o português e línguas do grupo banto, produziram alguns resultados semelhantes nos níveis fonológico, lexical e morfossintático. A pesquisadora defende um contínuo afro-brasileiro de português e, mostrando semelhanças compartilhadas pelas três variedades, acha difícil defender uma deriva natural da língua portuguesa.

Galves (2008) também traz alguns argumentos contra a hipótese da deriva e apresenta novos *corpora* com “(mais) pistas para uma nova agenda de pesquisa”. Segundo Galves, as variedades angolanas e moçambicanas podem nos ensinar muito sobre como as línguas africanas interferiram no desenvolvimento do português do Brasil. Comenta, por exemplo, das Atas da Sociedade dos Desvalidos de Salvador, fundada em 1832⁷, e de um conjunto de textos escritos em Angola, ao longo dos séculos 17, 18 e 19, do Arquivo Caculo Cahahenda. Tanto nas Atas escritas por africanos quanto nos textos escritos do Arquivo Caculo Cahahenda, encontramos muitos dos fenômenos variáveis presentes no português africano e no

⁷ “Dessas atas, algumas foram escritas por africanos, e outras, mais numerosas, por brasileiros.” (GALVES, 2008, p. 156).

português brasileiro de hoje. Dessa forma, Galves (2008) não acredita na deriva no desenvolvimento do PB e defende a transmissão linguística irregular.

A comparação das vertentes africanas e brasileiras do português, bem como a comparação de ambas com as línguas africanas com que estiveram em contato na sua história, vem mudar substancialmente a discussão ao trazer uma base empírica para o velho debate. Procurei mostrar que na balança empírica, esses dados são mais pesados e consistentes do que os dos defensores da deriva [...]

A transmissão irregular no quadro de uma aquisição de segunda língua com exposição insuficiente aos dados da língua-alvo, da qual temos uma imagem moderna em Moçambique e Angola hoje, parece dar conta do desenvolvimento histórico da variação encontrada no Brasil de hoje. (GALVES, 2008, p. 160-161).

Sabemos que a realização variável da concordância verbal no Brasil é um dos fenômenos que está no cerne das discussões sobre os processos que levaram à constituição do PB. Por esse motivo, estudos descritivo-comparativos como o nosso podem trazer novos elementos para o debate.

2 CONCORDÂNCIA VERBAL NO PB E NO PE: UMA REVISÃO

Com dados do português do Brasil, há vários estudos variacionistas voltados para a concordância verbal de 3ª pessoa do plural. Com dados de Portugal, o número de estudos é bem menor. Não é a nossa intenção apresentar todos os trabalhos que já foram realizados no Brasil. Iniciamos com o trabalho de Lemle e Naro (1977), considerado precursor nos moldes labovianos. Em seguida, trazemos alguns trabalhos representativos de diferentes variedades, com diferentes *corpora*, para uma visão geral do fenômeno em terras brasileiras. Nesta seção, não vamos descrever as variáveis linguísticas e sociais que os estudos controlaram, pois uma descrição mais detalhada da maioria dessas variáveis consta da metodologia deste trabalho.

Ao final do primeiro capítulo do livro “Doa-se lindos filhotes de poodle: variação linguística, mídia e preconceito” (2005), da sociolinguista Marta Scherre⁸, há uma ampla bibliografia sobre o assunto, que foi atualizada em 2008, na reimpressão do livro, tendo em vista que há muito feito sobre o tema, principalmente no PB. Remetemos à obra de Scherre para uma visão abrangente dessa literatura.

2.1 No Português do Brasil

Miriam Lemle e Anthony J. Naro (1977) desenvolveram um estudo da língua falada pelos estudantes do MOBRAL (Movimento Brasileiro de Alfabetização) do Rio de Janeiro, com vistas à verificação de pontos de discrepância ou de diferenciação entre a variedade de língua portuguesa utilizada por esse grupo social e a variedade de língua escrita de nível jornalístico e da literatura contemporânea mais acessível.

Porque na fala do grupo social a que pertenciam os estudantes do MOBRAL a concordância do verbo com o sujeito era um fenômeno variável, Lemle e Naro (1977) puderam utilizar o aparato teórico-metodológico da Teoria da Variação e Mudança Linguística que introduziu o conceito de regra variável.

O *corpus* foi composto por 20 informantes do Rio de Janeiro, 9 mulheres e 11 homens, numa faixa etária dos 17 aos 50 anos. Foram controladas três variáveis

⁸ A parceria de Marta Scherre e Anthony Naro rendeu inúmeros trabalhos sobre a concordância verbal (cf. NARO e SCHERRE, 1991, 1999a, 1999b, 2000, 2003a, 2003b, 2007, 2010; SCHERRE e NARO, 1993, 1998a, 1998b, 2000, 2005).

linguísticas (definição ou indefinição do sujeito, posição do sujeito em relação ao verbo e saliência fônica verbal) e uma extralinguística (variável estilística). Lemle e Naro (1977) não obtiveram os resultados esperados apenas para a variável estilística⁹.

Vale destacar que os autores concluíram que o estudo da regra de concordância verbal prova a necessidade de se introduzir, no modelo de funcionamento sincrônico da gramática, o conceito de saliência: “uma regra gramatical será mais ou menos aplicada, dependendo da saliência dos efeitos provocados” (LEMLE e NARO, 1977, p. 47).

O trabalho de Rodrigues (1987) sobre a concordância verbal de 3ª pessoa do plural foi realizado com dados de pessoas residentes em favelas da periferia da cidade de São Paulo. Rodrigues decidiu levar em conta a real composição populacional da capital paulistana, que abriga, principalmente na periferia, extenso contingente de migrantes, adultos de baixa ou nula escolaridade, procedentes da zona rural, não só do interior do estado de São Paulo, mas, fundamentalmente, de outras regiões do Brasil.

A pesquisa envolveu 40 pessoas adultas, caracterizadas em função do sexo (16 homens e 24 mulheres), da idade (20 a 35 anos, 36 a 50 anos, mais de 51 anos) e da escolaridade (analfabetos ou até a 4ª série do ensino fundamental). Dentre os fatores sociais, encontramos também o fator procedência dos informantes. Foram 1.356 dados analisados, com uma frequência de 71% de ausência de concordância verbal.

Os grupos de fatores gênero, escolaridade e procedência não se mostraram relevantes do ponto de vista estatístico para a explicação do processo. Em relação à procedência do informante, a sua hipótese foi de que os falantes paulistanos tenderiam a “errar” menos que falantes provenientes de outras regiões do Brasil. No entanto, a diferença entre os índices correspondentes à procedência dos informantes não foi tão notável, o que tornou possível afirmar que os brasileiros analfabetos ou de baixa escolaridade tendem, em geral, a não fazer concordância do verbo com o sujeito da 3ª pessoa do plural. Já o grupo de fatores idade se mostrou relevante e apontou os falantes mais velhos utilizando mais a variante padrão.

⁹ Com a variável estilística, Lemle e Naro (1977) controlaram o grau de formalidade do encontro. Realizaram 7 entrevistas de uma hora de duração para cada informante, em diversos locais.

Os três grupos de fatores linguísticos controlados por Rodrigues (1987) obtiveram relevância estatística (posição do sujeito em relação ao verbo, saliência fônica e presença/ausência do sujeito pronominal 'eles/elas').

Scherre e Naro (1993) apresentaram resultados sobre duas dimensões do paralelismo formal na concordância verbal: paralelismo formal no nível clausal (marcas no sujeito) e paralelismo formal no nível discursivo (marcas no verbo). Esse trabalho foi inédito em relação à proposta e avaliação das variáveis mencionadas na concordância verbal de 3ª pessoa do plural no PB. Os pesquisadores analisaram um total de 4.616 construções, que apresentaram um percentual de 73% (3.366/4.616) de concordância verbal. A amostra utilizada proveio do *Corpus* Censo do Programa de Estudos sobre o Uso da Língua (PEUL)¹⁰. Os resultados apresentados mostraram que marcas conduzem a marcas e zeros conduzem a zeros, tanto no nível clausal quanto no nível discursivo, evidenciando-se a tendência de formas gramaticais particulares ocorrerem juntas.

Vieira (1995) realizou um estudo variacionista sobre a concordância verbal na fala de pescadores norte-fluminenses (RJ). Os dados foram selecionados de 72 inquéritos do Arquivo Sonoro do Projeto APERJ (Atlas Etnolinguístico dos Pescadores do Estado do Rio de Janeiro). Todos os informantes são do sexo masculino, analfabetos ou com pouca escolaridade. A pesquisa mostrou uma alta produtividade do cancelamento da marca de plural (62%). Dos grupos de fatores de caráter linguístico, mostraram-se significativos a saliência fônica, o paralelismo nos níveis clausal e discursivo, e a posição do sujeito em relação ao verbo. As variáveis animacidade do sujeito e distância entre o núcleo do SN sujeito e o verbo apresentaram comportamento instável em relação à concordância verbal e demonstraram exercer influência de nível secundário.

Scherre e Naro (1998a) analisaram dois aspectos importantes: (1) o efeito do traço [humano] do sujeito sobre a concordância no português falado e (2) a interação entre o traço de número e o traço humano no controle da concordância em dados do português do Brasil escrito na década de 90 (português moderno) e em dados de documentos do português do século XIII ao XVI (português antigo). Os dados do português falado foram extraídos do *Corpus* Censo do PEUL. Os dados do

¹⁰ O *Corpus* Censo do Programa de Estudos sobre o Uso da Língua (PEUL) é constituído por 64 horas de fala, gravadas com falantes nascidos na cidade do Rio de Janeiro, subdivididos em função do sexo (32 mulheres e 32 homens), idade (de 7 a 71 anos) e anos de escolarização (de 1 a 11 anos).

português escrito moderno foram extraídos de revistas, de jornais de grande circulação, de dissertações de mestrado e tese de doutorado da área de comunicação e de letras, circulares, ofícios e memorandos do meio universitário, atas de condomínio de classe média, bulas, livros etc. E os dados do português antigo foram extraídos de um conjunto significativo de obras, abrangendo o período do século XIII ao século XVI: *A Demanda do Santo Graal* (século XIII), *Diálogos de São Gregório* (século XIV), *Boosco Deleitoso* (séculos XIV e XV), *Vida e Feitos de Júlio César* (século XV), *O Preste Ioam das Índias* (séculos XIV e XVI). Na análise desses três *corpora*, concluíram que o traço [humano] é significativo tanto no sentido de reter o controle da concordância quando o núcleo é [+ humano] quanto no sentido de deslocar o controle da concordância para o nome do SPrep [+ humano] plural.

No artigo “Influência de variáveis escalares na concordância verbal”, os resultados de Naro e Scherre (1999a) evidenciaram a relevância de três variáveis estruturais na variação da concordância verbal de 3ª pessoa do plural: (1) o número de sílabas que separa os sujeitos antepostos de seus respectivos verbos, (2) a saliência fônica na relação singular/plural e (3) o número de elementos contidos no sujeito plural. Os dados analisados foram extraídos do *Corpus Censo* do Programa de estudos sobre o Uso da Língua (PEUL). Os resultados mostraram uma diminuição progressiva de marca de plural no verbo à medida que aumenta o número de sílabas entre o sujeito e verbo. Em relação à saliência fônica, o estudo mostrou que o aumento da saliência do material fônico na oposição singular/plural dos verbos aumenta as chances da concordância verbal. Por fim, o aumento do número de elementos contidos no sujeito plural tende a diminuir as chances do uso do plural.

Monguillhott (2001) analisou a concordância verbal em uma amostra de língua falada pertencente ao Projeto Varsul (Variação Linguística Urbana na Região Sul) composta de 24 florianopolitanos estratificados de acordo com idade, sexo e escolaridade. De um total de 1.583 dados, 1.251 (79%) apresentaram a concordância verbal e 332 (21%) apresentaram a não concordância. Os resultados do estudo apontaram os fatores linguísticos (saliência fônica, posição do sujeito em relação ao verbo, paralelismo formal, traço semântico do sujeito, tipo de verbo e tipo de sujeito) como mais relevantes em detrimento dos fatores sociais (escolaridade e idade). No entanto, a escolaridade e a idade também obtiveram significância estatística.

A dissertação de Pereira (2004) também apresentou um estudo variacionista da regra de concordância verbal de 3ª pessoa do plural¹¹. O *corpus* foi composto de 15 inquiridos de informantes idosos (em média 78 anos) de ambos os sexos, analfabetos ou semiescolarizados, nascidos e criados na zona rural dos estados de São Paulo e Minas Gerais, na área correspondente às trilhas das bandeiras paulistas. Pereira (2004) analisou 520 dados de 3ª pessoa do plural e encontrou apenas 24% de aplicação da regra. Os grupos de fatores linguísticos selecionados conforme a relevância estatística foram: saliência fônica, presença/ausência do sujeito pronominal (eles, elas, vocês), paralelismo discursivo, traço semântico do sujeito, paralelismo oracional e papel semântico do sujeito. E os três grupos de fatores sociais controlados se mostraram estatisticamente significativos: gênero, procedência e escolaridade.

O trabalho de Gameiro (2005) com a língua falada contou com informantes de escolaridade nula a nível superior, procedentes de Araraquara e Itirapina, cidades da região central do Estado de São Paulo. Do total de 1.399 ocorrências de terceira pessoa do plural da amostra, 772 (55%) não trouxeram a marca formal de plural nos verbos, sendo que 627 (45%) apresentaram a marca formal de plural. Na ordem de relevância, as variáveis linguísticas e sociais selecionadas foram: escolaridade, saliência fônica, constituição morfossintática do sujeito, determinação do sujeito, paralelismo formal, categorização semântica do sujeito, transitividade verbal, gênero e idade.

Em nosso trabalho de Mestrado (MONTE, 2007), analisamos a variação na concordância verbal de 3ª pessoa do plural na fala de 20 pessoas residentes numa comunidade da periferia urbana da cidade de São Carlos, localizada no interior do Estado de São Paulo. As entrevistas sociolinguísticas informais foram feitas com jovens e adultos de 20 a 40 anos, de ambos os sexos, de procedência geográfica diversificada, diferenciados, também, com relação ao grau de escolarização: 10 não alfabetizados e 10 concluintes da oitava série do ensino fundamental na EJA (Educação de Jovens e Adultos).

Do total de 1.000 ocorrências estudadas no nosso *corpus*, 753 (75%) não apresentaram a marca formal de plural nos verbos, sendo que apenas 247 (25%) trouxeram a marca formal de plural. Dentre as variáveis controladas, a saliência

¹¹ Pereira (2004) também analisou a regra de concordância verbal de 1ª pessoa do plural.

fônica foi selecionada em primeiro lugar e apresentou o *range* mais alto (85). Em segundo lugar vem o paralelismo formal no nível oracional, com um *range* de 43. A terceira variável selecionada foi a presença/ausência do 'que' relativo, com o terceiro maior *range* (36). A escolaridade veio em quarto lugar (*range*: 20), e a última variável selecionada foi o gênero, apresentando o menor *range* (10). Os nossos resultados mostraram que os efeitos linguísticos foram mais fortes do que os efeitos sociais.

Scherre, Naro e Cardoso (2007) investigaram o efeito do tipo de verbo na concordância verbal no português brasileiro. Utilizaram no estudo três amostras distintas: (1) dados de 64 falantes da década de 80 (*Corpus Censo – PEUL*), (2) dados de 16 falantes do *Corpus Censo* recontactados em 1999/2000 e (3) dados de uma falante maranhense residente em Brasília (dados de Cardoso, 2005). Os autores realizaram um estudo minucioso com verbos categorizados numa perspectiva tradicional e também numa perspectiva gerativista. Concluíram que a única característica específica do verbo que influencia a concordância verbal é a saliência fônica da oposição singular/plural. Em momento algum o tipo de verbo, nem mesmo a classe dos inacusativos, revelou significância estatística.

Rubio (2008) investigou a concordância verbal de 3ª pessoa do plural na fala da Região Noroeste do Estado de São Paulo, mais precisamente na Região de São José do Rio Preto. Para a realização da sua pesquisa, Rubio constituiu uma subamostra, composta de 76 entrevistas, do Banco de Dados Iboruna¹². Do total de 3.308 ocorrências analisadas, 2.314 (70%) apresentaram marcas de plural explícitas nos verbos. A análise estatística realizada evidenciou as variáveis paralelismo formal no nível oracional, escolaridade, paralelismo formal no nível discursivo, saliência fônica, posição do núcleo do SN/sujeito em relação ao verbo, traço semântico do sujeito, idade, gênero e tipo morfológico do sujeito como as mais relevantes no condicionamento da variação.

O trabalho de Lucchesi, Baxter e Silva (2009)¹³ mostrou a variação do fenômeno em questão em três comunidades rurais afro-brasileiras isoladas do

¹² O Banco de Dados Iboruna foi composto pelo Projeto ALIP (Amostra Linguística do Interior Paulista), no período de março/2004 até setembro de 2007. Trata-se de iniciativa inédita, por constituir o primeiro banco de dados de amostras de fala do interior do Estado de São Paulo, com rigorosa coleta de dados e controle de fatores sociais, abrangendo sete municípios da região noroeste: Bady Bassitt, Cedral, Guapiaçu, Ipiguá, Mirassol, Onda Verde e São José do Rio Preto. (RUBIO, 2008, p. 62). O coordenador geral do projeto é o Prof. Dr. Sebastião Carlos Leite Gonçalves (GONÇALVES, 2003, 2007).

¹³ O trabalho é baseado nos resultados da pesquisa de Mestrado de Jorge Augusto Alves da Silva (2003), orientada por Dante Lucchesi.

interior do Estado da Bahia: Cinzento, Helvécia e as comunidades geminadas de Barra e Bananal. Os informantes eram todos analfabetos ou semianalfabetos. Foram depreendidas 1.706 ocorrências e a regra de concordância foi aplicada em 273, correspondendo a 16% do total. A análise variacionista revelou que seis fatores estruturais condicionam a aplicação da regra: saliência fônica, forma de indicação do plural no sujeito, concordância nominal no SN sujeito, realização e posição do sujeito, caracterização semântica do sujeito e tipo de verbo. Das variáveis extralinguísticas, três se mostraram estatisticamente relevantes: faixa etária, comunidade e sexo.

Em uma pesquisa posterior à de 2005, Gameiro (2009) analisou como o fenômeno se manifesta na língua escrita a partir de um *corpus* composto de redações elaboradas por alunos da Escola Estadual Zita de Godoy Camargo, situada na cidade de Rio Claro, no interior do estado de São Paulo. A pesquisadora trabalhou com redações de alunos de 5ª série (6º ano) do ensino fundamental até o 3º ano do ensino médio, inclusive com a EJA (Educação de Jovens e Adultos). Entretanto, a escolaridade não foi selecionada como fator relevante pelo programa Goldvarb. O índice de cancelamento de marca de plural nos verbos foi de 16%. Dentre os grupos de fatores controlados, o programa Goldvarb selecionou, na ordem de relevância, a saliência fônica verbal, o paralelismo discursivo, a posição do sujeito, a função sintática do elemento à direita do verbo, a presença/ausência do sujeito pronominal 'eles/elas', o paralelismo formal, o traço semântico do sujeito, o gênero e a presença/ausência do 'que' relativo.

Gameiro (2009) também aplicou um teste de percepção que revelou que o número de livros que os estudantes leem e a perspectiva de vida que eles têm influenciam na avaliação que fazem da concordância verbal. A pesquisadora concluiu que "os informantes que ambicionam profissões de médico e alto *status*, reconhecem com mais facilidade as construções pertencentes à norma popular ou padrão" (p.194).

O trabalho de Oliveira (2010) também é realizado na região noroeste do Estado de São Paulo. A autora analisou a variação da concordância verbal na língua falada e na língua escrita. Para a língua falada, utilizou 24 informantes do banco de dados IBORUNA e, para a língua escrita, mais de 600 redações escolares produzidas por alunos do 1º ano do ensino fundamental ao 3º do ensino médio. As redações escolares foram cedidas pela escola Profº Oscar Arantes Pires, escola

municipal de 1º ao 5º ano (ensino fundamental I) e pela escola estadual Profº Justino Jerry Faria, escola de 6º ano ao 3º colegial (ensino fundamental II e ensino médio). Ambas são escolas da cidade de São José do Rio Preto. Na modalidade falada, Oliveira encontrou 85% de presença de concordância verbal, com os seguintes grupos de fatores atuantes: paralelismo formal, distância e posição do sujeito, faixa etária, escolaridade, saliência fônica, estrutura do sujeito. Na modalidade escrita, o índice de frequência de concordância verbal também foi de 85% e os grupos de fatores selecionados foram: escolaridade, paralelismo formal, saliência gráfica e posição do sujeito.

No período em que realizamos as entrevistas sociolinguísticas para o nosso trabalho de Mestrado (MONTE, 2007), os nossos informantes concluintes do ensino fundamental na EJA (5 mulheres e 5 homens) também produziram vários textos dissertativos. Na ocasião, não foi possível analisar a produção escrita desses informantes. O trabalho foi concluído posteriormente (MONTE, 2011) e pudemos analisar e comparar a variação da concordância verbal de 3ª pessoa do plural na fala e na escrita das mesmas pessoas. Do total de 489 ocorrências estudadas na amostra de língua falada, 336 (68%) não trouxeram a marca formal de plural nos verbos, sendo que apenas 153 (31%) apresentaram a marca formal de plural. Na amostra de língua escrita, encontramos a concordância verbal predominando. Do total de 218 ocorrências, 59 (27%) não apresentaram a concordância verbal e 159 (72%) apresentaram a marca de plural nos verbos. Na modalidade falada, as variáveis saliência fônica, paralelismo formal no nível oracional e presença/ausência do 'que' relativo foram as mais relevantes. Já na modalidade escrita, apenas a variável saliência gráfico/fônica se mostrou estatisticamente significativa.

2.2 No Português Europeu

São poucos os trabalhos de variação na concordância verbal de terceira pessoa do plural no PE. Destacamos os seguintes estudos: Carrilho (2003), Varejão (2006), Naro e Scherre (2007), Monguilhott (2009), Gandra (2009), Almeida (2010), Bazenga (2011), Rubio (2012), Vieira (2012) e Brandão e Vieira (2012).

Carrilho (2003), na linha da linguística formal, revelou com dados do CORDIAL-SIN¹⁴ que a concordância sujeito-verbo na terceira pessoa do plural exhibe alguma variação. Segundo a pesquisadora, a “discordância” verbal surge, em variedades não padrão, em contextos em que o constituinte considerado sujeito se encontra em posição pós-verbal (construções com verbos inacusativos e construções predicativas).

Varejão (2006) também investigou a variação na concordância verbal de terceira pessoa do plural no PE com dados do CORDIAL-SIN. Para a constituição do *corpus*, a pesquisadora utilizou os registros transcritos que na ocasião estavam disponíveis na internet. Foram 2.520 ocorrências coletadas, sendo que 223 (9%) não apresentaram marcas de concordância. Varejão (2006) controlou as seguintes variáveis: saliência fônica da desinência verbal, tipo de verbo (‘ser’ *versus* outros tipos de verbos), forma verbal, posição do sujeito em relação ao verbo, estrutura do SN sujeito e estatuto sintático do verbo que carrega as marcas de concordância. A autora não realiza uma análise de regra variável rigorosamente variacionista. Por isso, não apresenta os resultados em termos de pesos relativos. Conclui que a saliência fônica e a posposição do sujeito são aspectos que levam à ocorrência de concordância não padrão também no PE popular.

Naro e Scherre (2007) coletaram e analisaram 235 dados de falta de concordância em oito textos do português medieval. Comprovaram, com pesos relativos, a influência das variáveis saliência fônica, posição do sujeito em relação ao verbo e traço semântico do sujeito. Segundo Naro e Scherre (2007, p. 65), “a concordância variável tanto no português europeu falado hoje quanto no português europeu medieval escrito exhibe as mesmas características estruturais fundamentais encontradas no português moderno falado no Brasil. As diferenças são uma questão de grau, não de tipo”.

Outro trabalho que investigou o fenômeno em questão no PE foi o de Monguilhott (2009). Com uma amostra constituída de dezesseis entrevistas gravadas na região de Lisboa, estratificada de acordo com idade e escolaridade, a pesquisa de Monguilhott (2009) revelou que o traço humano no sujeito, a posição do

¹⁴ O CORDIAL-SIN, *Corpus Dialectal para o Estudo da Sintaxe*, é um projeto do Centro de Linguística da Universidade de Lisboa que apresenta excertos selecionados de 4500 horas de gravações de fala realizadas em todo o território português. São falas de pessoas analfabetas ou de pouca escolarização.

sujeito em relação ao verbo e o tipo de verbo são fatores que condicionam a variação da concordância no português europeu. Dentre os fatores sociais, a escolaridade se mostrou relevante. As variáveis saliência fônica, paralelismo formal, tipo de sujeito, gênero e redes sociais (mobilidade e localismo) não foram selecionadas. Do total de 807 dados obtidos, 65 (8%) apresentaram a variante zero de plural nos verbos.

O estudo de Gandra (2009), assim como o de Carrilho (2003) e o de Varejão (2006), também se valeu de dados do ‘*Corpus Dialectal para o Estudo da Sintaxe*’ (CORDIAL-SIN). Foram analisadas 904 ocorrências de verbos com sujeito referencial na terceira pessoa do plural, e os resultados obtidos foram os seguintes: 32 ocorrências de perda de concordância (3,5%) e 872 ocorrências de marcação da concordância (96,5%). Das ocorrências sem concordância verbal, 14 ocorreram em contextos com o sujeito posposto ao verbo, 12 com o sujeito anteposto, 4 com sujeito retomado por um relativo e 2 com sujeito não realizado. Gandra (2009, p.153) destacou os dados com o verbo ‘ser’, dizendo que “nos casos com o verbo ‘ser’, é mais frequente a perda de concordância que no geral do *corpus*. Se o índice de perda de concordância é de 3,5% para o *corpus* como um todo, essa taxa sobe para 10% no uso do verbo ‘ser’”.

Almeida (2010) focalizou a concordância verbal de terceira pessoa do plural na modalidade escrita da variedade europeia, com base em produções de texto de estudantes do 9º ano do Ensino Básico e 12º ano do Ensino Secundário. As produções de texto foram coletadas em escolas da cidade de Lisboa. A amostra, composta por 1.733 dados, apresentou apenas 25 ocorrências (1,41%) sem a marca explícita de plural. Diante de poucos dados, não foi viável utilizar o programa Goldvarb-X para o tratamento estatístico. Das ocorrências sem concordância, 17 ocorreram com sujeito anteposto, 6 com sujeito posposto e 2 com sujeito nulo. O discreto índice de ausência de concordância permitiu à pesquisadora afirmar que se trata de uma regra semicategórica.

Bazenga (2011) analisou dados de 16 informantes do *Corpus MAD-FNC*¹⁵. Do total de 1.217 dados, 1.026 (85%) apresentaram a marca explícita de plural nos verbos e 191 (15%) apresentaram a marca zero de plural. As variáveis linguísticas

¹⁵ O *corpus* foi constituído por 21 pessoas residentes na cidade do Funchal, localizada na ilha da Madeira. Os informantes foram selecionados com base em três dimensões de estratificação social: sexo, faixa etária e nível de escolaridade.

significativas que favoreceram a ausência de marca de concordância foram: posição do sujeito, tipo semântico do sujeito e saliência fônica. A única variável social significativa foi o nível de escolaridade do informante. É importante informar que Bazenga incluiu formas verbais que, para os brasileiros, são homófonas (tem/têm, vem/vêm). Acreditamos que foram esses dados que elevaram o índice de ausência de concordância na cidade do Funchal (15%). Os trabalhos sociolinguísticos realizados por pesquisadores brasileiros excluem tais dados. Informamos, ainda, que os resultados apresentados por Bazenga (2011) são preliminares.

A investigação de Rubio (2012) contou com dados do *Corpus de Referência do Português Contemporâneo* (CRPC)¹⁶, do Centro de Linguística da Universidade de Lisboa. As 133 entrevistas utilizadas por Rubio (2012) foram retiradas do sub-corpus oral espontâneo do CRPC. Trata-se de entrevistas coletadas por pesquisadores portugueses, em diversas regiões de Portugal, entre as décadas de 1980 e 1990, com aproximadamente 10 minutos de duração. Do total de 1.107 ocorrências de terceira pessoa do plural estudadas, 1.039 (93,9%) trouxeram a marca formal de plural nos verbos, sendo que somente 68 (6,1%) não apresentaram a marca de plural. O programa estatístico Goldvarb selecionou apenas três variáveis linguísticas (posição do sujeito, traço semântico do sujeito e tipo morfológico do sujeito), das seis consideradas. Das variáveis sociais investigadas (escolaridade, faixa etária e gênero), nenhuma foi selecionada.

Os trabalhos de Vieira (2012) e Brandão e Vieira (2012) foram desenvolvidos no âmbito do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português* (www.letras.ufrj.br/concordancia). Em Portugal, foram realizadas 18 gravações em Lisboa/Oeiras e 18 gravações em Cacém, cidade-dormitório vizinha de Lisboa, consoante as variáveis idade, escolaridade e sexo. Vieira (2012) analisou 1.477 construções de 3ª pessoa do plural das gravações de Lisboa/Oeiras e a amostra apresentou apenas 19 dados de não-concordância (1,3%). Brandão e Vieira (2012) analisaram 1.515 ocorrências da comunidade de Cacém e encontraram apenas 17 dados sem a marca de plural nos verbos (1,1%). Com poucos dados de ausência de concordância verbal, foi inviável o uso do programa de análise estatística

¹⁶ O CRPC é um vasto *corpus* da variedade europeia do Português e de outras variedades (Brasil, Angola, Cabo Verde, Guiné-Bissau, Moçambique, São Tomé e Príncipe, Goa, Macau, Timor-Leste). Contendo 311,4 milhões de palavras, este *corpus* abrange diferentes tipos de textos escritos (literário, jornalístico, técnico, etc.) e de registros orais (formal e informal).

Goldvarb-X. Por isso, as pesquisadoras realizaram uma análise qualitativa das ocorrências encontradas. Os dados dos dois estudos não apresentaram comportamento diferente, nem quantitativa nem qualitativamente. Destacaram a ausência de concordância verbal em estruturas com o verbo 'ser' e em construções com sujeito posposto. Segundo as autoras, em termos quantitativos e qualitativos, não é possível argumentar em prol das semelhanças entre PB e PE quanto ao fenômeno.

3 O UNIVERSO DESTE ESTUDO

Nesta seção, apresentamos o universo de nossa pesquisa: as comunidades de fala onde o fenômeno foi estudado e quais foram os procedimentos metodológicos adotados para que a investigação fosse desenvolvida. Dessa forma, organizamos a presente seção em duas partes, cada uma contemplando um dos aspectos referidos acima.

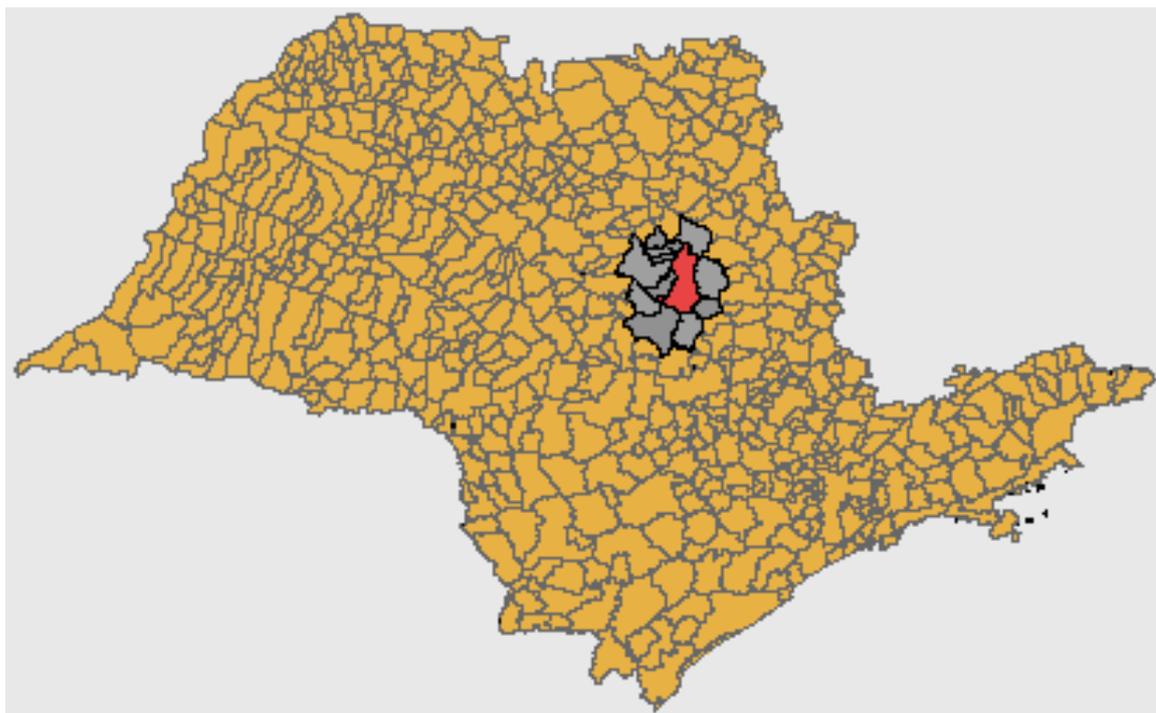
3.1 As comunidades de fala

A pesquisa foi realizada com a língua falada da cidade de São Carlos, localizada no interior do Estado de São Paulo, região Sudeste do Brasil, e com a língua falada da cidade de Évora, situada no Alentejo, sul de Portugal.

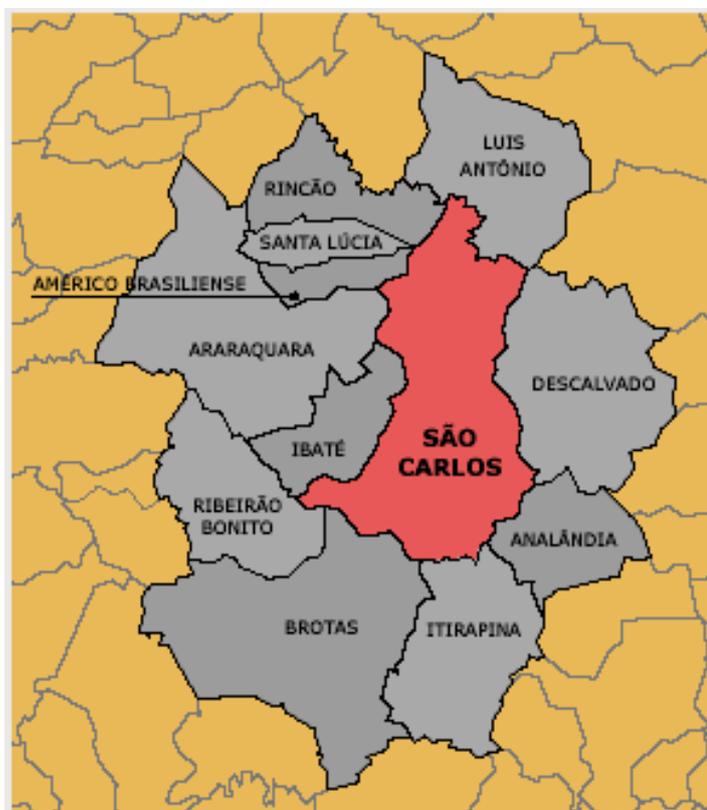
Escolhemos a cidade de Évora, em Portugal, por apresentar algumas características que se aproximam da cidade de São Carlos, no Brasil. Assim como São Carlos, Évora é uma cidade interiorana e universitária. Em termos populacionais, as duas são consideradas cidades de médio porte.

Apresentamos, a seguir, breves considerações sobre a cidade de São Carlos e sobre a cidade de Évora.

3.1.1 Considerações sobre a cidade de São Carlos



Mapa 1 – Localização do Município de São Carlos no Estado de São Paulo
Fonte: São Carlos (2005)



Mapa 2 – São Carlos e municípios vizinhos
Fonte: São Carlos (2005)

A cidade de São Carlos está localizada no centro geográfico do Estado de São Paulo¹⁷. De acordo com o Censo 2010 do IBGE¹⁸, a população de São Carlos é de 221.950 habitantes, sendo 213.061 na área urbana e 8.889 na área rural. São 108.914 homens e 113.036 mulheres.

A cidade destaca-se como centro regional tecnológico e industrial. A atividade universitária é intensa e devido à presença, principalmente, de duas instituições públicas de ensino superior estabelecidas na cidade, a Universidade de São Paulo (USP) e a Universidade Federal de São Carlos (UFSCar), é grande a atividade de pesquisa e a concentração de cientistas no município.

Segundo o *site* oficial da prefeitura, São Carlos possui um pesquisador doutor para cada 180 habitantes. No Brasil, a relação é de um doutor para cada 5.423 habitantes¹⁹. A edição 35/2011 da revista Kappa magazine (edição especial – 154 anos de São Carlos) traz uma reportagem atualizando o número de doutores na cidade²⁰. Segundo o mais recente levantamento realizado pelo Prof. Dr. Jorge Oishi, docente e assessor estatístico da reitoria da UFSCar, São Carlos conta com um doutor para cada 140 habitantes. Nesse levantamento, entram apenas os 1.600 doutores das cinco instituições ligadas ao ensino e pesquisa – USP, UFSCar, UNICEP e das duas unidades da Embrapa.

O Censo 2010 do IBGE revela que a taxa de analfabetismo em São Carlos é de 3,46%. A cidade possui 6.751 pessoas não alfabetizadas de 10 anos ou mais de idade.

Quanto à atividade industrial, ainda segundo dados obtidos junto à prefeitura, destaca-se a presença de indústrias de motores, compressores, lápis, geladeiras e fogões, além de empresas têxteis e uma grande quantidade de indústrias de médio e pequeno porte, de diversos setores de produção. Já na agropecuária, é relevante a produção de leite, cana-de-açúcar, laranja, frango, carne bovina e milho.

¹⁷ Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), o Estado de São Paulo conta com 645 municípios. A população do Estado de São Paulo, segundo o Censo 2010, é de 41,2 milhões de habitantes. No Brasil, os primeiros resultados definitivos, divulgados em novembro de 2010, apontaram uma população formada por 190.732.694 pessoas. Esses dados estão disponíveis em: www.ibge.gov.br.

¹⁸ Dados obtidos do *site* www.ibge.gov.br.

¹⁹ Dados obtidos do *site* www.saocarlos.sp.gov.br.

²⁰ KAPPA MAGAZINE. São Carlos: Abelhaneda Editora, ano 2, n. 35, 24 out. 2011. Edição especial: 154 anos de São Carlos.

De acordo com a Fundação Pró-Memória²¹, o histórico de São Carlos revela que os seus primeiros habitantes foram os indígenas, principalmente dos grupos tupi e jê. A região era conhecida antigamente como “Campos de Araraquara” ou “Sertões de Araraquara”. Os primeiros não indígenas a ocupar a região foram os bandeirantes, que exploravam o interior do Brasil em busca de ouro, pedras preciosas e índios para o trabalho escravo. A existência oficial da cidade começou com a demarcação das primeiras sesmarias da região. Foi em 1785 que se demarcou a primeira, a Sesmaria do Pinhal. Em 1810, foi demarcada a do Monjolinho e, por último, em 1812, a do Quilombo.

No ano de 1856, começou a ser erguida uma capela, por iniciativa de alguns proprietários de terra, especialmente herdeiros da família Arruda Botelho, dando origem ao processo de fundação da cidade. Em 4 de novembro de 1857, a imagem de São Carlos Borromeu foi trazida da sede da Fazenda Pinhal e colocada no altar da capela. É neste dia que se comemora tanto o dia do santo quanto a fundação oficial da cidade²².

Entre os anos de 1831 e 1857, são formadas as lavouras cafeeiras, de modo que o café passou a ser a principal atividade econômica de São Carlos, tornando-se, inclusive, produto de exportação. A chegada da ferrovia, em 1884, contribuiu para que o produto fosse facilmente escoado para o porto de Santos, firmando a importância da cidade no cenário econômico e político. Quem trabalhava nas fazendas cafeeiras eram negros escravizados que vinham de outras regiões do Brasil, principalmente da Bahia, uma vez que o tráfico de escravos vindos da África foi abolido em 1850.

A sociedade dessa época era composta por um pequeno grupo que fazia parte da elite, os fazendeiros e seus familiares, e uma grande maioria da população era de escravos. Alguns trabalhadores livres, que geralmente faziam serviços nas fazendas, formavam uma pequena parcela da população. Apenas com a chegada da ferrovia, o comércio foi impulsionado e diversos profissionais, como médicos, professores, advogados, etc., passaram a compor a classe média urbana.

Truzzi (2007) traz uma análise das contribuições étnicas, espontâneas ou não, determinantes para a estrutura de composição da população que se formou

²¹ A Fundação Pró-Memória, criada em 1993, mantém e disponibiliza os documentos dos poderes executivo, legislativo e judiciário, objetivando também a preservação do patrimônio artístico e arquitetônico do município.

²² No dia 4 de novembro de 2012, comemoramos os 155 anos da cidade de São Carlos.

sobre o território. Sejam formações espontâneas ou não, é necessário ressaltar que os grupos populacionais estabelecidos foram essenciais para a manutenção do sistema agrícola, sendo importante considerar que

toda presença ou extermínio, marginalização ou absorção dos grupos populacionais em questão encontra sua razão de ser na implantação e sobrevivência do sistema de grandes lavouras inaugurado em meados do século XIX no oeste paulista. Matança dos índios, expulsão dos posseiros, importação dos negros escravos e depois dos imigrantes europeus são etapas de um mesmo processo que se sobrepõem, uma após a outra. (TRUZZI, 2007, p.47).

Os escravos, como já mencionamos, contribuíram fortemente na manutenção da produção cafeeira, sendo trazidos, sobretudo, de províncias nordestinas após a abolição do tráfico negreiro. A tabela 1 apresenta o número de escravos existentes em São Carlos e em Araraquara entre os anos de 1874 e 1887, e revela que o emprego de escravos praticamente não diminuiu até a abolição:

Tabela 1: População escrava em São Carlos e Araraquara, de 1874 a 1887.

	1874	1883	1884	1885	1886	1887
São Carlos	1.568	3.465	3.774	3.773	2.982	3.726
Araraquara	1.626	2.247	1.828	1.811	1.300	1.627

Fonte: (TRUZZI, 2007, p. 53).

São Carlos, em 1874, já possuía quase a mesma quantidade de escravos que Araraquara. Vale destacar que até 1887, ano anterior à abolição, o número de escravos praticamente não diminuiu, excetuando-se no ano de 1886 em que foi promulgada a lei dos Sexagenários. Vê-se ainda que, em vários anos, São Carlos apresentou quase o dobro do número de escravos em comparação ao de Araraquara.

São Carlos foi, de fato, um grande ponto de comércio de escravos. Com base no Censo Parochial de 1874, Rizzoli (1995) menciona que “a população negra do município, computando os escravos, pardos livres e pretos livres, correspondia a 39,3%. A população do município era de 7.897 indivíduos, destes 1.568 escravos, sendo 926 do sexo masculino e 642 do sexo feminino”.

Próximo à abolição, a mão de obra assalariada já estava sendo implantada por alguns fazendeiros, pois acreditavam que o trabalho dos imigrantes poderia preservar o sistema de grandes lavouras. Após a abolição, os escravos não tiveram perspectivas de inserção e mobilidade social, vivendo marcados pela marginalização e discriminação. A maioria deles, que vivia na Fazenda Conde do Pinhal, por exemplo, saiu da zona rural e aglutinou-se na periferia da cidade. Há registros de que eles formaram um bairro de negros, existente ainda hoje e conhecido por Vila Isabel (cf. TRUZZI, 2007, p. 63).

Segundo Rizzoli (1995), a exclusão do negro do mercado de trabalho livre o levou à exclusão social. Dessa forma, o negro passou a ser visto como “problema” para a sociedade. O autor transcreve algumas notas policiais veiculadas pela imprensa local de São Carlos, exemplificando tal situação, entre elas a que segue, extraída do Correio de São Carlos de 01 de janeiro de 1900:

Aproveitamos da ocasião que essa notícia oferece para lembrar o Dr. Delegado de Polícia que seria bom organizar patrulhas mesmo de dia pelos subúrbios da cidade onde há vendas em que constantemente se vê magotes de vagabundos. Esses vagabundos são em geral pretos e pretos que ali passam o dia em completa ociosidade dando lugar ao vício da embriaguez. (RIZZOLI, 1995).

Nas últimas décadas do século XIX e no início do século XX, ocorreu o fenômeno social que deixou grande influência nos hábitos, costumes e cultura da região: a imigração. Os imigrantes vieram para trabalhar nas lavouras de café e, graças às suas habilidades, atuaram também na manufatura, no comércio e contribuíram para a formação da cidade. De acordo com Truzzi (2007, p. 65-67), durante mais de 20 anos, até 1904, a grande maioria dos imigrantes chegados em São Carlos eram provenientes da Itália. No final do século XIX, a cidade de São Carlos se constituía um dos polos mais atrativos de imigração do Estado de São Paulo. O pesquisador apresenta os seguintes dados da imigração na cidade:

Tabela 2: Total de imigrantes em São Carlos e posição em relação ao Estado de São Paulo entre 1884 e 1903²³.

Ano	Italianos	Portugueses	Espanhóis	Outros	Total	Posição em relação ao Estado de São Paulo
1884	–	–	–	–	304	3º
1886	458	76	17	2	553	3º
1887	–	–	–	–	926	4º
1894	–	–	–	–	3.788	1º
1895	–	–	–	–	4.444	2º
1898	–	–	–	–	1.342	4º
1900	–	–	–	–	869	4º
1901	2.822	114	459	137	3.532	4º
1902	1.381	82	–	24	1.487	3º
1903	257	18	32	1	308	5º

Fonte: (TRUZZI, 2007, p. 66)

Segundo o autor, em 1886, cerca de um oitavo (2.051 habitantes) da população total do município (16.104 habitantes) era de imigrantes. Entre esses, mais da metade era de italianos (1.050 indivíduos); em seguida, vinham os portugueses (464).

Depois da década de 1930, o café não era mais a principal atividade econômica do município. São Carlos, que já era uma cidade importante na época, viu crescer uma série de outras atividades econômicas, iniciadas pelos imigrantes que traziam de seus países técnicas diversas e montavam seus estabelecimentos comerciais. Assim, era grande o número de sapateiros, alfaiates, fabricantes de massas e de outros gêneros de consumo diário. A cidade ganhou novos serviços, teve suas ruas pavimentadas e, com a crise cafeeira, muitas pessoas deixaram a zona rural, fazendo com que a cidade crescesse consideravelmente.

O número de indústrias na cidade aumentou após a 2ª Guerra Mundial, acompanhando a tendência nacional. Nos anos de 1950, a cidade foi pioneira na produção de geladeiras em série. Também nessa época foi criada a Escola de Engenharia na cidade – a USP – e, na década de 70, foi fundada a Universidade Federal de São Carlos (UFSCar). A cidade, nessa época, atraía muitas indústrias e estudantes, aumentando o número de moradores do município. Hoje, a cidade que já passa dos 200 mil habitantes, é reconhecida pelo seu polo científico e recebeu

²³ Para alguns anos, Truzzi (2007) não obteve a procedência dos imigrantes.

recentemente, da presidenta Dilma Rousseff, pela lei aprovada em 13 de outubro de 2011, o título de “Capital da Tecnologia”.



Fazenda Pinhal – São Carlos
Fonte: arquivo pessoal



Catedral de São Carlos
Fonte: arquivo pessoal



Estação Cultura – São Carlos
Fonte: arquivo pessoal



Escola Estadual Álvaro Guião – São Carlos
Fonte: arquivo pessoal

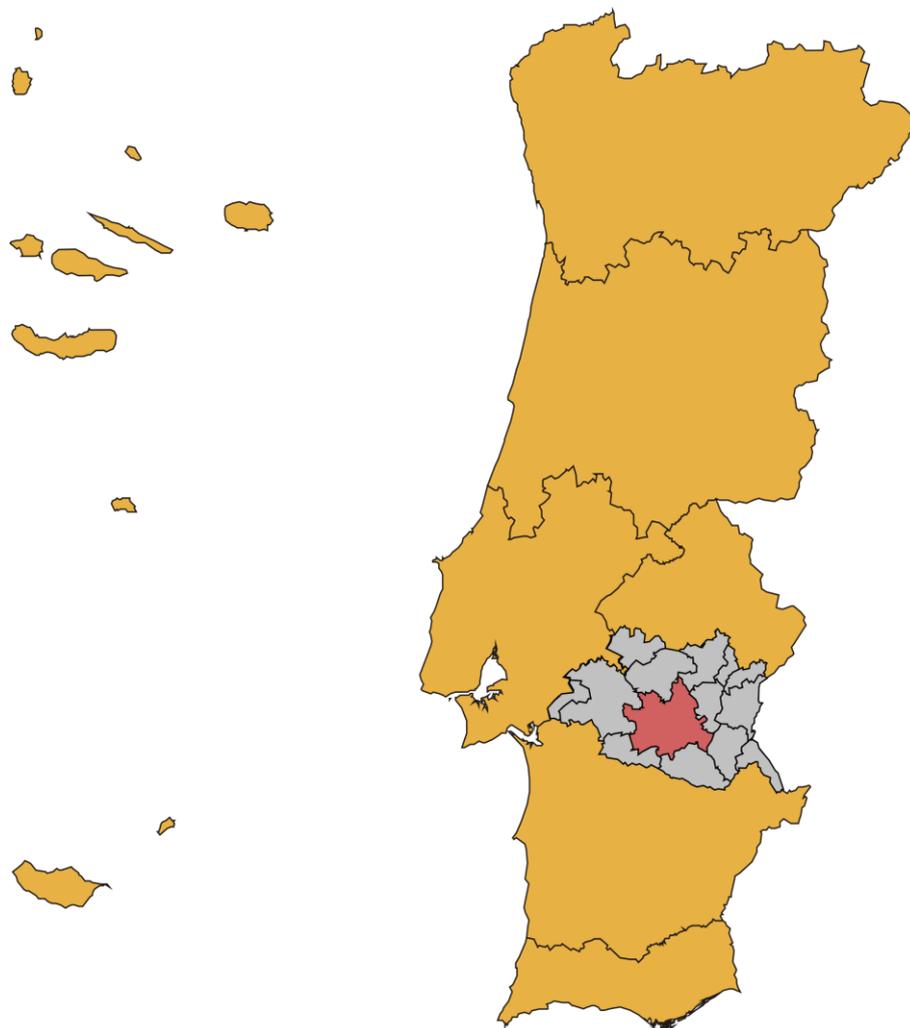


Fazenda Santa Maria – São Carlos
Fonte: arquivo pessoal



Universidade Federal de São Carlos
Fonte: arquivo pessoal

3.1.2 Considerações sobre a cidade de Évora²⁴



Mapa 3 – Localização do Município de Évora em Portugal²⁵
Fonte: INE (Instituto Nacional de Estatística)

²⁴ A maioria das informações aqui apresentadas se encontra no site www.cm-evora.pt da Câmara Municipal de Évora e no trabalho de Simplício (1997).

²⁵ Os resultados definitivos do Censo 2011 indicam que a população residente em Portugal é de 10.562.178 habitantes. Na última década, verificou-se um abrandamento do crescimento demográfico, com 2% de aumento, face aos 5% observados na década de 90. Esses dados estão disponíveis em: www.ine.pt.



Mapa 4 – Localização do Município de Évora na região do Alentejo
Fonte: INE (Instituto Nacional de Estatística)



Mapa 5 – Évora e municípios vizinhos
Fonte: INE (Instituto Nacional de Estatística)

Évora é uma cidade branca como uma ermida. Convergem para ela os caminhos da planície como o rasto da esperança dos homens. E como a uma ermida, o que a habita é o silêncio dos séculos, do descampado em redor. Conheço, dos seus espectros, a vertigem das eras, a noite medieval mora ainda nas ruas que se escondem pelos cantos, nas pedras cor do tempo ouço um atropelo de vozes seculares. Vozes de população, gritos de condenados, ecos de reis, senhores, estrépito de guerras, ódios e sonhos, sob a imobilidade dos mesmos astros. Como um cofre do tempo, irrealizado e absoluto, a cidade ignora a exactidão do presente, conhece apenas o alarme da memória. As casas novas têm todas a mesma idade de séculos. E quando se sai da cidade, a planície prolonga, até a um limite irreal, esta voz de infinitude.

Vergílio Ferreira

O Concelho de Évora integra-se numa vasta planície que se estende ao sul de Portugal – a região do Alentejo²⁶. De acordo com os resultados definitivos do Censo 2011 do Instituto Nacional de Estatística (INE), a população residente de Évora é de 56.596 habitantes, sendo 26.831 homens e 29.765 mulheres.

A cidade de Évora é o principal polo urbano da região, em termos populacionais e funcionais. Segundo informações do *site* da Câmara Municipal, a dinâmica social e econômica da cidade tem conseguido contrariar a tendência da região no seu conjunto, mantendo um crescimento idêntico ao de outras cidades médias portuguesas. O turismo tem-se afirmado como uma das áreas mais relevantes das atividades econômicas de Évora. Diversas atividades complementares (comércio, construção, etc.) são significativas no município, devido à presença de um mercado de consumo de dimensão apreciável, o qual tem sido beneficiado pela presença da Universidade de Évora.

Em relação à época da fundação da cidade, diz Simplício:

Apesar do interesse dedicado à clarificação das questões relacionadas com as origens da cidade de Évora, não foi ainda possível obter uma opinião de consenso dos diversos investigadores, no que se refere à época e iniciativa da fundação da cidade. (SIMPLÍCIO, 1997, p. 87).

²⁶ A população residente no Alentejo, segundo os resultados definitivos do Censo 2011, é de 757.302 habitantes. Em termos regionais, a evolução demográfica da última década indica que a região do Alentejo volta a perder população, registrando uma diminuição de 19.283 pessoas face a 2001. É no Alentejo que se verifica a maior taxa de analfabetismo, com 9,6%. Os resultados definitivos do Censo 2011 estão disponíveis no *site* do Instituto Nacional de Estatística (www.ine.pt).

De qualquer modo, não há dúvida de que Évora tem origem muito remota. A região de Évora é um território antigo, com vestígios de ocupação humana desde a pré-história, de que sobreviveram gravuras rupestres e importantes monumentos megalíticos, como a Anta do Zambujeiro e o Cromesque dos Almendres.

Posteriormente, Évora foi ocupada pelos Romanos, tendo sofrido forte romanização numa época de grande importância para a cidade. A posição geográfica privilegiada pode explicar, em parte, a importância que a cidade adquiriu desde a antiguidade. Durante o período romano, Évora adquire o estatuto de município de direito latino e a designação honorífica de *Liberalitas Julia* do Imperador Júlio César. Este período caracteriza-se por um desenvolvimento notável, tanto sob o ponto de vista socioeconómico, quanto artístico e monumental.

Segundo Simplício (1997, p. 91), após a queda do Império Romano, Évora vai sofrer um longo período de domínio visigótico e árabe, durante o qual decaiu bastante a sua importância cultural, apesar de se manter um centro económico e militar importante. Desses períodos pouco restou, mas, segundo Alexandre Herculano, talvez baseado nos escritos de um geógrafo árabe do séc. XII, Évora seria então a segunda cidade árabe da província de Badajoz e grande entreposto comercial, com uma produção agrícola diversificada.

A ocupação árabe terminou em 1165, quando Giraldo Sem Pavor conquistou a cidade e a integrou na coroa de D. Afonso Henriques. Évora iria conhecer um novo período de desenvolvimento e reforçar a sua posição na rede urbana do país. A conquista de Évora aos mouros por Giraldo Sem Pavor, em 1165, marcou uma nova fase de crescimento da cidade até o séc. XVI, quando atingiu o seu máximo esplendor, na sequência da fixação da Corte e do processo histórico dos Descobrimentos.

A dinastia de Avis teve certamente um papel fundamental no desenvolvimento da cidade: D. João I eleva Évora a segunda cidade do reino (no fim do século XV, teria cerca de 10.000 habitantes), e a presença da Corte, por períodos prolongados, faz deslocar para Évora, além das importantes famílias portuguesas, uma plêiade de artistas que marcaram em definitivo a fisionomia da cidade. É nesta fase que se verifica a instalação do Palácio Real junto ao Convento de S. Francisco, que a Universidade de Évora é fundada e são construídos os mais belos palácios, conventos e igrejas e o Aqueduto da Água da Prata, obra fundamental que provocou algumas alterações na rede urbana.

Em 1759, com a reforma do sistema universitário e com a expulsão dos Jesuítas do país, na sequência da política pombalina, a Universidade de Évora é extinta e só será restaurada daí a outros dois séculos.

Posteriormente a toda a atividade renovadora associada ao século XIX e princípios do século XX, que visava dotar a cidade das infraestruturas necessárias à melhoria das condições de vida, o sucesso da ditadura salazarista, a partir da década de 20, traduziu-se no abrandamento significativo do desenvolvimento e a cidade entrou numa fase de declínio e apagamento que duraria até a década de 70, em que a Revolução do 25 de abril de 1974 (Revolução dos Cravos) inverteu este processo.

Tendo servido de palco a inúmeros povos que por ela passaram e nela se fixaram, uns por curtos períodos, outros por longos séculos, Évora espelha diversos estilos e tendências estéticas. Conta com uma reconhecida qualidade arquitetónica, onde podemos admirar decorações em azulejo e trabalhos em ferro forjado que remontam aos séculos XV a XVIII.

O centro histórico, orgulho lusitano, foi classificado pela UNESCO como Patrimônio da Humanidade em 1986. A cidade assume claramente uma vocação patrimonial, cultural e universitária.



Praça do Giraldo – Évora
Fonte: arquivo pessoal



Templo Romano e Biblioteca Pública – Évora
Fonte: arquivo pessoal



Templo Romano – Évora
Fonte: arquivo pessoal



Catedral de Évora
Fonte: arquivo pessoal



Praça do Giraldo – Évora
Fonte: arquivo pessoal



Universidade de Évora
Fonte: arquivo pessoal



Largo da Porta de Moura – Évora
Fonte: arquivo pessoal



Muralha de Évora
Fonte: arquivo pessoal

3.2 Procedimentos metodológicos da pesquisa

Em todo empreendimento acadêmico que lide com pesquisa na comunidade de fala, existe sempre muito interesse quanto aos primeiros passos a dar. A pergunta: “O que você diz às pessoas?” não é trivial.” (LABOV, 2008 [1972], p. 242)²⁷.

Nesta segunda parte da seção, mais técnica, apresentamos os *corpora*, a maneira como foi a coleta e a seleção dos dados, bem como os grupos de fatores linguísticos e sociais.

3.2.1 Os *corpora* sob análise

Constituímos dois *corpora* semelhantes entre a cidade de São Carlos/SP (Brasil) e a cidade de Évora (Portugal). É importante ressaltar que não utilizamos o *corpus* da cidade de São Carlos do nosso trabalho anterior (cf. MONTE, 2007), porque os informantes eram de procedência geográfica diversificada. Sendo assim, um novo *corpus* da comunidade de fala são-carlense foi constituído. A coleta em Évora foi realizada durante o estágio de Doutorado no exterior de fevereiro a dezembro de 2010²⁸.

Cada *corpus* é composto de 18 entrevistas de aproximadamente uma hora de duração entre informante e documentador (DID)²⁹. Trabalhamos com seis células, sendo cada célula formada de três informantes, estratificados em função do gênero e da escolaridade³⁰. São três homens e três mulheres não alfabetizados, três

²⁷ In any academic course that deals with research in the speech community, there is always a great deal of interest in the first steps to be taken: “What do you say to people?” This is not a trivial question. (LABOV, 1972, p. 207).

²⁸ O estágio de Doutorado no exterior ocorreu no período de fevereiro a dezembro de 2010, na Universidade de Évora, sob a supervisão da Profa. Dra. Maria Filomena Candeias Gonçalves. O estágio foi realizado com bolsa CAPES, processo 4614/09-7, no âmbito do Programa de Doutorado no País com Estágio no Exterior (PDEE).

²⁹ A nomenclatura vem do *corpus* do Projeto de Estudo da Norma Urbana Linguística Culta (NURC). Esse *corpus* compõe-se de entrevistas gravadas durante as décadas de 1960/70 nas cinco cidades brasileiras que tinham então mais de um milhão de habitantes: Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre. Para documentar essa variedade, foi entrevistado um total de 2.356 pessoas, em 1.570 horas de gravação. As entrevistas, todas com pessoas de formação universitária, seguem três formatos diferentes: elocuições formais – geralmente aulas universitárias –, diálogos entre o informante e o documentador e diálogos entre dois informantes, com a participação marginal de um documentador.

³⁰ A idade dos informantes não foi considerada uma variável independente, pois cada nível de escolaridade corresponde a uma determinada faixa etária. Os estudantes concluintes do ensino médio/secundário, por exemplo, estavam todos entre 17 e 18 anos.

homens e três mulheres que estavam terminando o ensino fundamental/básico na EJA (Educação de Jovens e Adultos) e três homens e três mulheres que estavam terminando o ensino médio/secundário regular. Vale informar que o ensino básico em Portugal corresponde ao ensino fundamental no Brasil. Já o ensino secundário corresponde ao ensino médio.

Abaixo, temos a distribuição dos informantes segundo escolaridade e gênero nas duas comunidades estudadas:

São Carlos (BRASIL)			
	Homens	Mulheres	Total
Não alfabetizados	3	3	6
Ensino Fundamental – EJA	3	3	6
Ensino Médio	3	3	6
Total	9	9	18

Évora (PORTUGAL)			
	Homens	Mulheres	Total
Não alfabetizados	3	3	6
Ensino Básico – EJA	3	3	6
Ensino Secundário	3	3	6
Total	9	9	18

Os quadros que apresentamos a seguir resumem as características dos nossos informantes de São Carlos (PB) e de Évora (PE):

Informantes de São Carlos (PB)			
Informante	Gênero	Idade	Escolaridade
Z	F	54 anos	não alfabetizada
I	F	66 anos	não alfabetizada
P	F	72 anos	não alfabetizada
T	M	38 anos	não alfabetizado
B	M	58 anos	não alfabetizado
R	M	59 anos	não alfabetizado
S	F	36 anos	Ensino Fundamental – EJA
U	F	42 anos	Ensino Fundamental – EJA
E	F	44 anos	Ensino Fundamental – EJA
C	M	31 anos	Ensino Fundamental – EJA
L	M	45 anos	Ensino Fundamental – EJA
O	M	46 anos	Ensino Fundamental – EJA
A	F	17 anos	Ensino Médio
J	F	17 anos	Ensino Médio
M	F	18 anos	Ensino Médio
V	M	17 anos	Ensino Médio
H	M	17anos	Ensino Médio
G	M	18 anos	Ensino Médio

Informantes de Évora (PE)			
Informante	Gênero	Idade	Escolaridade
M	F	79 anos	não alfabetizada
B	F	80 anos	não alfabetizada
G	F	91 anos	não alfabetizada
W	M	71 anos	não alfabetizado
J	M	78 anos	não alfabetizado
R	M	84 anos	não alfabetizado
N	F	40 anos	Ensino Básico – EJA
L	F	42 anos	Ensino Básico – EJA
C	F	46 anos	Ensino Básico – EJA
P	M	23 anos	Ensino Básico – EJA
U	M	33 anos	Ensino Básico – EJA
F	M	53 anos	Ensino Básico – EJA
E	F	17 anos	Ensino Secundário
I	F	17 anos	Ensino Secundário
T	F	18 anos	Ensino Secundário
O	M	17 anos	Ensino Secundário
H	M	17 anos	Ensino Secundário
D	M	18 anos	Ensino Secundário

3.2.2 A coleta dos dados

Os passos elementares de localizar e contactar informantes e levá-los a falar livremente numa entrevista gravada são problemas difíceis para os estudantes. É um erro ignorar essas questões, pois nas práticas e técnicas que têm sido elaboradas estão contidos vários princípios importantes do comportamento linguístico e social. (LABOV, 2008 [1972], p. 242)³¹.

As entrevistas sociolinguísticas informais com os informantes, tanto em São Carlos quanto em Évora, foram realizadas por nós de modo que se aproximassem da língua falada do dia a dia. De acordo com Rodrigues (1987), o problema que se coloca ao pesquisador que pretende recolher amostras do vernáculo de qualquer comunidade a que ele não pertence é exatamente o da dificuldade de se transformar num “igual linguístico” com relação ao seu informante. Essa dificuldade foi expressa por Labov, por meio do que ele denominou “paradoxo do observador”:

[...] o objetivo da pesquisa linguística na comunidade deve ser descobrir como as pessoas falam quando não estão sendo sistematicamente observadas – no entanto, só podemos obter tais dados por meio da observação sistemática. (LABOV, 2008 [1972], p.244)³².

Se o objetivo da pesquisa no seio da comunidade é descobrir como as pessoas falam quando não são observadas sistematicamente, a mera presença do entrevistador com o seu gravador pode afetar o estilo de fala do informante. Assim, o pesquisador terá de agir com cautela, para reduzir ao máximo os efeitos do chamado paradoxo do observador.

Para Labov (1972, p. 85-86), os contextos variam quanto ao nível de formalidade: informal e formal. Já os estilos podem ser de três tipos, dependendo do contexto: se o contexto for informal, o estilo de fala será casual; se o contexto for formal, o estilo de fala poderá ser monitorado ou espontâneo.

³¹ The elementary steps of locating and contacting informants, and getting them to talk freely in a recorded interview, are formidable problems for students. It is an error for anyone to pass over these questions, for in the practices and techniques that have been worked out are embodied many important principles of linguistic and social behavior. (LABOV, 1972, p. 207).

³² [...] the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation. (LABOV, 1972, p. 209).

Contexto:	Informal	Formal
Estilo:	Casual	Monitorado/Espontâneo

Segundo Labov, o *corpus* ideal para o estudo linguístico é a fala utilizada em situações cotidianas. Este constitui o que Labov denomina estilo casual, ou seja, a fala do dia a dia das pessoas, usada em situações informais, em que a atenção não está voltada para a linguagem.

No entanto, toda observação sistemática de um falante, como é a situação de uma entrevista, condiciona um contexto formal, em que o grau de atenção voltado ao discurso é maior. Por isso, diversas técnicas foram utilizadas com o objetivo de suscitar o estilo casual na situação de entrevista.

Resumimos, assim, os nossos dois desafios: minimizar ao máximo os efeitos do chamado paradoxo do observador e, dessa forma, conseguir um estilo casual na situação de entrevista. Para isso, no decorrer das nossas leituras, recolhemos um bom número de conselhos a respeito de como agir frente à comunidade.

Adotamos vários procedimentos, sendo que muitos deles foram explicitados por Rodrigues (1987), Berlinck (1988), Tarallo (2002), pelos pesquisadores do Projeto Filologia Bandeirante (1998 *apud* Pereira, 2004) e Lucchesi (2009). Todos se basearam em Labov (1972):

- a) não dissemos ao informante que se tratava de uma pesquisa sobre língua, mas sim de uma pesquisa sobre aspectos sociais e culturais da comunidade;
- b) as entrevistas aconteceram na própria comunidade: nas residências, nos centros comunitários ou na escola ali existente;
- c) não seguimos um roteiro fixo de perguntas, mas partimos das informações já colhidas na fase do preenchimento da ficha social. Dessa forma, orientamos o diálogo para temas de interesse do informante, ligados à sua rotina de vida, obtendo assim um conjunto significativo de narrativas pessoais.
- d) buscamos nos acomodar à linguagem do informante, minimizando as diferenças de nossa fala em relação à do entrevistado.

Antes de iniciarmos a gravação da entrevista, como já mencionamos, preenchíamos uma ficha com as características sociais do informante. Esses

primeiros momentos foram de extrema importância, pois procurávamos deixar o informante mais à vontade, num clima bem descontraído e de confiança.

Em Portugal, os primeiros dias foram essenciais para conhecer a comunidade, os costumes e as tradições. Os informantes mais idosos, por exemplo, gostavam muito de explicar como era a vida antes e como ficou depois do 25 de abril de 1974 (Revolução dos Cravos). Os mais jovens conversavam sobre o sistema educativo português, expectativa de vida, a escolha do curso superior, a crise econômica na Europa, etc. As touradas, uma tradição ainda forte e polêmica no Alentejo, e a Feira de S. João, uma feira que acontece todo ano na cidade de Évora, renderam boas narrativas. Os portugueses, assim como os brasileiros, gostam muito de futebol e novela. Esses temas sempre suscitavam boas conversas informais. Outro tema produtivo com os informantes mais velhos, nas duas comunidades, foi a vida no campo, porque muitos moraram e trabalharam na zona rural.

De acordo com Lucchesi (2009, p.160), “o leque de temas numa entrevista de tipo sociolinguístico é potencialmente ilimitado, não havendo em princípio tema ruim”. O sucesso na obtenção da informalidade nas conversas é evidenciado nas longas histórias contadas pelos informantes. Os assuntos foram diversos: histórias da comunidade, narrativas populares, infância (as brincadeiras), religião, festas (religiosas ou não), namoro, casamento, família, desigualdade social, política, planos de vida, projetos de trabalho, sonhos, lazer, acesso a serviços médicos e à escola, etc.

Enfim, fazemos das palavras de Tarallo (2002) as nossas:

Os módulos cobrem uma série de tópicos para fins de conversação: dados pessoais do informante (sua história), jogos e brincadeiras de infância, brigas, namoro e encontros amorosos, casamento, perigo de morte, medo, família, religião, amigos, turmas, serviços públicos, o crime nas ruas, escola e trabalho, interação com outros membros da comunidade, esportes etc. **O sucesso da aplicação dos módulos poderá variar para cada comunidade de fala, para cada grupo de falantes ou mesmo para cada indivíduo. Cabe, portanto, ao investigador adaptá-los a cada grupo estudado!** (TARALLO, 2002, p. 22, grifo nosso).

Não podemos deixar de concordar com Tarallo (2002, p. 32) quando afirma que “quanto mais tempo você passar no campo, coletando dados, mais criativo você se tornará em relação às possíveis maneiras de minimizar o efeito negativo causado por sua participação direta na interação”.

3.2.3 A seleção dos dados

Não consideramos todas as ocorrências encontradas nos *corpora*. Resolvemos considerar as primeiras oitenta ocorrências de cada entrevista, a fim de obtermos um número mais ou menos equivalente para todos os informantes. No *corpus* do PE, conseguimos as oitenta ocorrências de cada entrevista. No *corpus* do PB, em quatro entrevistas não obtivemos os oitenta dados³³. Dessa forma, analisamos 1.440 ocorrências na amostra do PE e 1.422 ocorrências na amostra do PB.

A presente investigação contempla o estudo com sujeitos/SNs de estrutura simples e complexa de 3ª pessoa do plural representados por:

a) Nome substantivo no singular com um ou mais determinantes no plural.

- (1) os professor **tá** falando comigo eu tô viajando (M2PB)³⁴
- (2) os dois menino dela **são** esquizofrênico (F1PB)
- (3) os agricultor já não **davam** trabalho a ninguém... (M0PE)
- (4) e os lavrador **morava** lá e o meu pai era lá empregado (F0PE)

b) Nome substantivo no plural.

- (5) as palestras assim **são** perfeitas... (F2PB)
- (6) as novelas não **deviam** de ser assim... (M1PE)
- (7) as pessoas não **tão** conversando bem comigo (M1PB)
- (8) mas já há aí economistas que **dizem**... é melhor sairmos do euro... (F2PE)

³³ Informante Z (não alfabetizada – 54 anos) = 74 dados
 Informante B (não alfabetizado – 58 anos) = 74 dados
 Informante R (não alfabetizado – 59 anos) = 76 dados
 Informante H (ensino médio – 17 anos) = 78 dados

³⁴ A codificação que segue os exemplos indica as seguintes informações sobre o/a informante: a primeira letra refere-se ao sexo (**F** – feminino ou **M** – masculino); a segunda à escolaridade (**0** – não alfabetizado/a, **1** – formação supletiva (EJA) ou **2** – ensino médio/secundário) e as duas últimas à variedade do português (**PB** – português brasileiro ou **PE** – português europeu).

c) Pronomes pessoais ‘eles / elas’ com referência determinada.

- (9) telefonei pra lá... pa *os meus amigos* e eles **contactaram** com a empresa (M1PE)
- (10) olha *os meus netos...* **começaram** eles com doze anos (F0PE)
- (11) *meus irmão trabalhô* lá... não sei que setor que eles **trabalhava** (M1PB)
- (12) *mai a minhas prima* tamém **mora** aí/ aí... acho que doi quarterão daqui... pra cima da escolinha ali... elas também **aprenderu** (F0PB)

d) Pronome pessoal ‘eles’ com referência indeterminada.

- (13) apesar de ser inglês de décimo segundo aquilo que eles **pedem...** (F2PE)
- (14) desta assembleia que eles **andam** a arranjar (M0PE)
- (15) a água era poço que eles **fazia... fazia** na rua poço (M1PB)
- (16) agora que eles **fizeram** uma pista de skate ali... (F2PB)

e) Pronome ‘vocês’. Ainda que ‘vocês’ se refira do ponto de vista semântico e discursivo à 2ª pessoa do plural, sob uma perspectiva gramatical tal pronome exige verbos com marca formal de 3ª pessoa do plural, o que justifica a sua inclusão no estudo da concordância de 3ª pessoa.

- (17) cêis não **vão** consegui (um) bom emprego (F2PB)
- (18) vocês não **prestam** pra nada (M0PE)
- (19) nem com um animal cêis não **pode** fazê isso (M1PB)
- (20) porque vocês **tratam** muito (por) você... (M1PE)

f) Outros pronomes.

- (21) muitos **morre** po/com o calor agora nesse tempo... (M0PE)
- (22) então dá pa contá os que **tá** lá faz tempo (F1PB)

- (23) otas **faziam** malha... (F1PE)
- (24) porque acho que as otra já (ti:)... sei lá se **recebeu** ou se não **recebeu**... (M0PB)
- (25) acho que só alguns é que **conseguem** mesmo... (M2PE)
- (26) depois da consulta que todos **foram** embora eu fiquei pro final (F2PB)

g) Numeral no plural.

- (27) duas ou três **vão** fazer faculdade (M2PB)
- (28) seis milhões **são** benfiquistas... (M2PE)
- (29) quatro não... **morreu** cinco... (M0PB)
- (30) dez **fumam**... é a maioria... (F1PE)

h) Dois ou mais núcleos (sujeito composto).

- (31) comunidade e escola **vai** sê se Deus quisé um dia uma coisa só né? (F1PB)
- (32) porque o nome em si e a escritura... **eram** deles não é? (M1PE)
- (33) o meu sogro e minha sogra tamém **veio/ veio** da Itália... (F0PB)
- (34) o meu pai os meus tios e o meu avô **vieram** um ano... numa fera (F1PE)

Além dos casos enumerados acima (sujeitos/SNs explícitos), foram incluídos os casos de sujeito nulo com referente no trecho do discurso.

i) Sujeito nulo.

- (35) (há) aí *aquelas pessoas* que **vivem**... ... Ø **vive** a maneira deles... (M1PE)
- (36) **passou** *uns cara de moto* e Ø **atirô** nele... (F2PB)

- (37) eu tenho *algumas amigas minhas* que **estão** nesse curso... que é um curso ahn:: direcionado... pronto... pra/ pra desporto pra atividade física... ahn *elas saem* de lá e já \emptyset **podem/ já podem** ir trabalhar como técnicas mesmo em ginásios... (F2PE)
- (38) *eles batem* mesmo... \emptyset já:: **tentaru** me matá... né?... foi por Deus que eu não morri... \emptyset **levaru** eu... num determinado lugar... no/ no/ num rio... \emptyset **bateru** a noite intera... e::: \emptyset me **afogaru**... \emptyset **deru** choque tudo... aí \emptyset **dexaru** eu... desmaiado lá... (M1PB)

Vale dizer que também incluímos o verbo 'ser' em estruturas clivadas, pseudoclivadas, em orações com valor existencial e nas designações de tempo, distância, quantidade, valor.

j) Verbo 'ser' em estruturas clivadas e pseudoclivadas.

- (39) **são** pessoas que gosta(m) da comunidade... (F1PB)
- (40) é que normalmente **era** as pessoas que **tinham** lá famílias a trabalhá já... (M0PE)
- (41) **foi** muitas coisas que me/ que/ que **feiz** eu desanimá... (F1PB)
- (42) quem compra estes porcos pretos cá em Portugale... **são** os espanhóis... (M1PE)
- (43) quem trabaivava nessa parte **era** só os home... (M0PB)
- (44) agora quem tá no nosso governo **são** os socialistas... (F2PE)

k) Verbo 'ser' em orações com valor existencial.

- (45) era assim... **era** sempre muitas pessoas... (M1PE)
- (46) **era** aquelas injeção (F0PB)
- (47) **é** as tendinites... (M1PE)
- (48) éh:: MMORPG... **seriam** jogos de RPG on line... (M2PB)

I) Orações com o verbo 'ser' nas designações de tempo, distância, quantidade, valor – casos em que as gramáticas prescrevem a concordância com o predicativo no plural.

- (49) era quatro hora... quatro e meia... treis e meia (F0PB)
- (50) parece que era quatro marcos... era quatro marcos a hora... (M0PE)
- (51) são cinco aulas... (F2PB)
- (52) são dez milhões... (M1PE)

3.2.4 Critérios de exclusão dos dados

O presente estudo não contempla todos os dados. Os critérios de exclusão aqui adotados são os seguintes:

a) Formas verbais que no singular e no plural não se distinguem na pronúncia por serem homófonas: *tem/têm*, *vem/vêm*. O mesmo critério foi utilizado nos dados do PE.

- (53) famílias pobres... têm/tem todas as condições (M1PE)
- (54) os moço vêm/vem conversá co Cido (F0PB)
- (55) hoje eles têm/tem mais liberdade... (M1PB)
- (56) porque eles vêm/vem da Ásia... (F2PE)

b) Respostas em que se repete a forma verbal da pergunta feita pelo documentador.

- (57) Doc.: jogam bem?
Inf.: jogam... dos grandes clubes jogam bem (F1PE)
- (58) Doc.: eles permitem a entrada?
Inf.: algumas vez(es) permitem... permitem... (M2PB)
- (59) Doc.: aqui... os jovens... na sua idade... podem começá a trabalhá ou não?
Inf.: podem... a partir dos dezesseis já podem () em part times... (M2PE)
- (60) Doc.: porque essas são da Globo né?
Inf.: são todas da Globo... (F2PB)

c) Verbo no plural com casa vazia do sujeito indeterminado (sem referente ‘eles’ no trecho do discurso) – contexto em que a marca de plural do verbo seria condicionada não pelo critério sintático (concordância com o sujeito/SN), mas pelo critério semântico (noção de indeterminação do sujeito).

- (61) meu vô que fazia... fez todo meu enxoval... calça blusinha **robaru** tudo... **entrou** dentro de casa **robaru** tudo... (F2PB)
- (62) cá também **jogam** muito com isso... (M1PE)
- (63) **mataru** ele... **mataru** ele mai num... num foi a muié não... (M0PB)
- (64) porque **acabaram** com o cinema que nós tínhamos (F2PE)

d) Orações com verbo ‘ter’ com valor existencial. Segundo as lições da gramática tradicional, o verbo ‘ter’ não deve ser usado no sentido de ‘haver’ (existencial). Entretanto, a construção é muito usual no PB.

- (65) e **tinha** aquelas parreira de... de/ de uva... (F0PB)
- (66) vixi... **tinha** vários doce lá... (M0PB)
- (67) **teve** doze pessoas só da minha sala... (F2PB)

e) Sujeito representado pelo pronome indefinido ‘tudo’ remetendo a um SN de 3ª pessoa do plural.

- (68) o meus irmão **tudo sabe** um poco... (F0PB)
- (69) e os filhos e o marido **é tudo** Benfica... (F1PE)
- (70) **são tudo** casado (M1PB)
- (71) esses desportes **é tudo** obrigatório na(s) escolas... (M2PE)

f) Sujeito representado por substantivo coletivo no singular que pode desencadear a chamada concordância semântica.

- (72) o time deles **são** muito entrosado (M2PB)
- (73) e o povo pobre **vivia** nos campos... (M1PE)
- (74) o pessoal **tirava** água do poço (M1PB)
- (75) pa o pessoal podê usá... aí **começar** usá... (M2PB)

g) Sujeito constituído por expressão partitiva. Segundo a tradição gramatical, quando o sujeito é constituído por expressão partitiva e um substantivo ou pronome plural, o verbo pode ir para o singular ou para o plural. 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 (CUNHA e CINTRA, 2001, p. 499).

- (76) a maioria tá indo com medo... **acha** que não **vai** conseguir (F1PB)
- (77) aliás... a maioria dos assaltos não **é** cometido por ciganos... (F2PE)
- (78) a maioria dos meus amigos assim **vão** pa barzinho (F2PB)
- (79) a maior parte deles **são** de aldeias (F1PE)

h) Verbos no infinitivo pessoal. De acordo com Brandão e Vieira (2012, p. 1058), “a forma infinitiva constitui por si só um contexto absolutamente particular, motivo pelo qual muitos estudos variacionistas nem sequer a consideram na contagem dos dados”.

- (80) porque não havia pessoas para **trabalhare** (M0PE)
- (81) pagou aos pais dele para **dizerem** que ele tinha morrido... (F2PE)
- (82) já temos/ já temos aí países que há mulheres a **jogarem** futebol (M1PE)
- (83) colocaram esses/ essas duas entidades pra me **atormentá**... (F2PB)

i) Sujeito composto quando está posposto ao verbo e com o núcleo mais próximo no singular. De acordo com as gramáticas normativas (cf. ROCHA LIMA, 1998; CUNHA e CINTRA, 2001), se o verbo vier antes do sujeito composto, ele pode ir para o plural ou concordar com o núcleo que estiver mais próximo.

- (84) **morreu** a senhora e o patrão... (F0PE)
- (85) **tava** o::: marido e a mulher... (M1PB)
- (86) na liga europa... **tá** o Porto e o Sporting... (M1PE)
- (87) ele queria que:::... eu e minha irmã ficasse por lá... que **voltasse** minha mãe meu pai pra cá... (M2PB)

j) Pronomes interrogativos. A princípio, estávamos considerando os casos com pronome interrogativo, mas na amostra do PE apareceram 7 ocorrências, e na amostra do PB apenas 1 ocorrência. Não sendo um número significativo quantitativamente, achamos melhor excluir esses dados.

- (88) também tem mais doentes... não sei quantos **são** aí... (F0PE)
- (89) a gente não sabe... quantos **são**... porque a gente vai... (F0PE)
- (90) e a gente não sabe quantos **são**... (F0PE)
- (91) é Brasil Portugal quais é que **são** as otras equipas? (M2PE)
- (92) posso ver quais **são** as que me posso candidatar... (F2PE)
- (93) quais é que **são** as vossas expectativas que tem pa o futuro... (F1PE)
- (94) e ele dava-nos uns dossiês... onde tinha quais é que **era** os objetivos... (F1PE)

- (95) as oitavas eu não sei quantas **são**... (F2PB)

3.2.5 Grupos de fatores (ou variáveis)

Nossa variável dependente é binária, ou seja, se constitui de duas variantes: presença de concordância verbal ou ausência de concordância verbal.

Tendo em vista o pressuposto teórico de que a variação linguística não é aleatória, é fundamental identificar conjuntos de circunstâncias linguísticas e sociais que tendem a favorecer ou desfavorecer o uso de uma ou outra variante. Esses conjuntos de circunstâncias linguísticas e sociais são denominados 'grupos de fatores'.

Um grupo de fatores representa uma das variáveis independentes, seja ela linguística ou social, que o pesquisador quer testar como uma possível influência no comportamento da variável dependente. (GUY e ZILLES, 2007, p. 238).

Muitos estudos analisados foram fundamentais para a definição dos grupos de fatores linguísticos e sociais, como também para várias das nossas hipóteses (cf. LEMLE e NARO, 1977; NARO e LEMLE, 1977; NARO, 1981; BORTONI-RICARDO, 1981; RODRIGUES, 1987; 1989, 2000; NARO e SCHERRE, 1991, 1999a, 1999b, 2000, 2003a, 2003b, 2007, 2010; SCHERRE e NARO, 1993, 1998a, 1998b, 2000, 2005; SCHERRE, NARO e CARDOSO, 2007; VIEIRA, 1995, 2012; MONGUILHOTT, 2001, 2009; CARRILHO, 2003; PEREIRA, 2004; GAMEIRO, 2005, 2009; VAREJÃO, 2006; MONTE, 2007; RUBIO, 2008, 2012; MOTA e VIEIRA, 2008; GANDRA, 2009; LUCCHESI, BAXTER e SILVA, 2009; ALMEIDA, 2010; OLIVEIRA, 2010; BAZENGA, 2011; BRANDÃO e VIEIRA, 2012).

Abaixo, segue a descrição dos grupos de fatores linguísticos e sociais estabelecidos para a análise da variação da concordância verbal da terceira pessoa do plural no PB e no PE.

3.2.5.1 Grupos de fatores linguísticos (ou variáveis linguísticas)

GRUPOS DE FATORES	FATORES
1) Grau de saliência fônica da oposição entre as formas verbais do singular e do plural	– 1a (<i>vive/vivem</i>) – 1b (<i>fala/falam</i>) – 1c (<i>quer/querem</i>) – 2a (<i>vai/vão</i>) – 2b (<i>bateu/bateram</i>) – 2c (<i>teve/tiveram</i>)
2) Paralelismo formal no nível oracional	– forma de plural explícita no último (ou único) elemento – forma zero de plural no último elemento – forma de plural explícita no último elemento de um SPrep – forma zero de plural no último elemento de um SPrep – presença de numeral no último elemento
3) Posição do sujeito/SN em relação ao verbo	– sujeito/SN anteposto – sujeito/SN posposto
4) Traço semântico do sujeito/SN	– SN [+ humano / + animado] – SN [– humano / + animado] – SN [– humano / – animado]
5) Distância entre o sujeito/SN e o verbo em termos do número de sílabas	– 0 a 2 sílabas – 3 a 5 sílabas – 6 a mais sílabas
6) Presença ou ausência do ‘que’ relativo ou complementizador	– ausência do ‘que’ – presença do ‘que’
7) Tipo estrutural do sujeito/SN	– pronome ‘vocês’ – pronome pessoal ‘eles/elas’ – SN pleno simples – SN pleno nu – SN pleno composto – numeral (como núcleo ou adjunto) – quantificador indefinido – pronome demonstrativo – sujeito nulo
8) Tipo de verbo: verbo ‘ser’ versus outros verbos	– verbo ‘ser’ – outros verbos

► Grau de saliência fônica da oposição entre as formas verbais do singular e do plural

O princípio da saliência fônica foi proposto por Lemle e Naro (1977) nos primeiros estudos sociolinguísticos sobre o tema. No nosso trabalho, estamos utilizando a escala de saliência fônica que foi reformulada por Naro em 1981, tendo em vista que muitos dos trabalhos desenvolvidos após essa data utilizam a escala reformulada, que pode ser vista em Naro e Scherre (1999a).

A escala compreende dois níveis (conforme a intensidade dos segmentos fonéticos que realizam a oposição), e seis classes (conforme a crescente diferença material entre as formas verbais do singular e do plural):

NÍVEL 1 (oposição não marcada / menos saliente): contém os pares cujos segmentos fonéticos que estabelecem a oposição são não marcados (são inacentuados e estão em sílaba átona] em ambos os membros.

1a: envolve nasalização sem mudança na qualidade da vogal na forma plural (*conhece/conhecem, vive/vivem, sabe/sabem*)

1b: envolve mudança na qualidade da vogal na forma plural (*ganha/ganham, era/eram, gosta/gostam*)

1c: envolve acréscimo de segmento na forma plural (*faz/fazem; quer/querem; diz/dizem; traz/trazem*)

NÍVEL 2 (oposição marcada / mais saliente): contém os pares cujos segmentos fonéticos que estabelecem a oposição são marcados (são acentuados) em pelo menos um membro da oposição.

2a: envolve ditongação e/ou mudança na qualidade da vogal na forma plural (*está/estão; dá/dão; vai/vão*)

2b: envolve acréscimo de segmentos sem mudanças vocálicas na forma plural (*viu/viram, foi/foram, bateu/bateram*)

2c: envolve acréscimos de segmentos e mudanças diversas na forma plural: mudanças vocálicas na desinência, mudanças na raiz, e até mudanças completas (*veio/vieram; disse/disseram; é/são*)

A expectativa do efeito da saliência fônica foi estabelecida por Lemle e Naro (1977) e, assim como os pesquisadores, acreditamos que quanto maior for a diferença entre as formas verbais do singular e do plural, maior será a probabilidade de realização da concordância e, por outro lado, quanto menor essa diferença, menor a chance de realização da concordância.

► **Paralelismo formal no nível oracional (marcas do sujeito)**

Esse grupo de fatores se baseia no estudo de Scherre e Naro (1993). Em nossa pesquisa, estamos analisando o paralelismo formal no nível oracional (marcas no sujeito), a fim de verificar se sujeitos com marcas explícitas favorecem mais a ocorrência de verbos com marcas explícitas de plural. Esse grupo de fatores abrange cinco categorias, ilustradas a seguir:

a) **presença da forma de plural explícita no último (ou único) elemento não inserido em um sintagma preposicional (S_{Prep})**

- (96) eles batem mesmo (M1PB)
- (97) os lavradoreses que **tomaram** depois conta das herdades (M1PE)
- (98) pessoas boass que me **ajudaru** a saí dessa situação... (F1PB)
- (99) havia unss que **davam** os tais dossiês a outros (F1PE)

b) **presença da forma zero de plural no último elemento não inserido em um sintagma preposicional (S_{Prep})**

- (100) minhas filha_ vai na igreja (M1PB)
- (101) e a mãe dela falô que os oto_ ia ficá falando... (F1PB)
- (102) os filho_ não quis mai morá lá naquela casa lá (M0PB)
- (103) e os lavrador_ morava lá e o meu pai era lá empregado... (F0PE)

c) presença da forma de plural explícita no último elemento inserido em um sintagma preposicional (SPrep)

- (104) são coisas de miúdos_u que **acontecem** acho com toda a gente né?...
(F2PE)
- (105) os montes de dívidas e montes de problemas_u que **surgiram** na firma
(F1PE)

d) presença da forma zero (plural ou singular) no último elemento inserido em um sintagma preposicional (SPrep)

- (106) nem as irmã da igreja_ **sabe** essas coisa (F1PB)
- (107) as irmãs da igreja_ **vai** lá (F1PB)

e) presença de numeral no último elemento

- (108) quatro ou cinco que **vai** lá fazê (M1PB)
- (109) os oto dois **estudam** (M0PE)
- (110) há três que **é** a filosofia... o português e a educação física... é a ginástica... (F2PB)
- (111) os três **vão** prestá Letras... (F2PB)

Estamos postulando a mesma hipótese de Scherre e Naro (1993), ou seja, “marcas levam a marcas e zeros levam a zeros”. Esperamos mais concordância verbal quando os sujeitos apresentarem marcas de plural explícitas no último (ou único) elemento.

Scherre e Naro (1993) são levados a concluir que o princípio da economia linguística não pode ser evocado para explicar o comportamento deste e de outros diversos fenômenos do português, bem como os de outras línguas naturais, com relação à influência da variável em questão.

O comportamento dos sintagmas por nós analisados constitui, portanto, evidência adicional para se levantar a hipótese a respeito da existência de um novo princípio de natureza universal associado ao funcionamento da variável paralelismo formal nas línguas naturais, ao lado de outros universais linguísticos já conhecidos. (SCHERRE e NARO, 1993, p.12).

Labov (1996), tendo conhecimento dos resultados obtidos por Scherre e Naro, afirma:

Los resultados de Scherre y Naro no dejan dudas de que la constricción dominante sobre la variación en el marcado de los verbos portugueses no es de orden funcional. Desde el punto de vista del planteamiento funcional tradicional, lo que estamos viendo no es una tendencia a preservar información semántica sobre la superficie, sino más bien una tendencia a emplear marcas donde menos se necesitan, y a omitirlas donde más se necesitan. Lo que no es decir que este paralelismo de estructura quede sin su contribución a la economía del lenguaje. Más bien, inscribe el estudio de la variación en las flexiones al estudio general de la concordancia en el lenguaje, área donde se ha reconocido largamente que los planteamientos funcionales son muy difíciles de sostener. (LABOV, 1996, p. 864).

É importante informar que há trabalhos (cf. PEREIRA, 2004) que questionam a posição de Labov segundo a qual o paralelismo seria de natureza contrafuncional. Sobre esse debate, Scherre (1998, p. 46) vai dizer que “considerar ou não a variável paralelismo como funcionalista ou não-funcionalista vai depender da amplitude do sentido da palavra funcionalista”.

► Posição do sujeito/SN em relação ao verbo

Os pioneiros no estabelecimento dessa variável foram Lemle e Naro (1977). Os pesquisadores postularam o princípio da saliência posicional, ou seja, as relações sintáticas entre sujeito/SN e verbo ficam mais salientes quando o sujeito/SN determinante e controlador da concordância verbal antecede o verbo determinado e candidato à concordância.

a) sujeito/SN anteposto:

- (112) houve lá pessoas que **arranjaram** lá muito dinheiro... (M0PE)
- (113) na época que eles **faleceru** eu era novinho... (M0PB)
- (114) e muitos ciganos **são** pobres... (F2PE)
- (115) acho que:: essas matérias assim **entram** no lugar da educação física né? (F2PB)

b) sujeito/SN posposto:

- (116) também às vezes lá **morava** peessoas... (F0PE)
 (117) na Santa Felícia **morria** três... dois por noite... (M1PB)
 (118) onde **tá** os campos de futebol ()... (M2PE)
 (119) aí **desce** as entidades nos médiuns pra podê limpá o terrero... (F2PB)

É muito comum encontrarmos nos trabalhos a conclusão de que o SN posposto ao verbo passa a ser encarado como objeto pelo falante, que não realiza a concordância. As primeiras pesquisadoras que apresentaram essa ideia foram Decat (1983)³⁵ e Eunice Pontes (1986). Vejamos:

De fato, poderíamos dizer que aqueles SNs, ao serem pospostos aos verbos das sentenças, perdem, entre outras, a propriedade sintática de sujeito, qual seja, a posição de início de sentença. Ainda que ocorram casos de o verbo concordar com o SN que está depois dele, isso não constitui evidência de que esse SN é sujeito... (DECAT, 1983, p. 17).

A conclusão, a respeito do chamado sujeito posposto, é que ele não tem as características que tipificam o sujeito em português, a não ser a concordância verbal, no registro escrito, formal, que é imposta pela gramática e o ensino escolar. Verificamos que o sujeito típico e o “sujeito posposto” são dois conjuntos disjuntivos. O único traço em comum é o da CV, traço por si muito frágil, porque ensinado nas escolas é praticamente inexistente na maioria dos registros.

Deve-se continuar considerando esse SN como sujeito? Minha conclusão é que ele não é sentido pelos falantes como sujeito. Por isso, eles não fazem a concordância.

Se o SN posposto não é sujeito, o que é? Se examinarmos os objetos diretos em português, veremos que o SN posposto tem mais traços de objeto do que de sujeito. Em primeiro lugar, a posição pós-verbal. Além disso, semanticamente, está mais para paciente do que para agente: tipicamente inanimado, indefinido. (PONTES, 1986, p. 172-173).

Berlinck (1988) analisou três *corpora*, sendo dois históricos (cartas de cunho pessoal do século XVIII (1750) e do século XIX (1850)) e um sincrônico, de 1987, que foi composto a partir de conversações gravadas com 20 jovens (20 a 30 anos), naturais de Curitiba/PR, de nível universitário. O trabalho de Berlinck revela uma frequência gradualmente menor de SNs pospostos à medida que se passa de um

³⁵ A autora deixa claro que a questão de se saber se o SN posposto continua, ou não, sujeito estava fora dos propósitos do seu trabalho.

momento a outro, em direção ao *corpus* sincrônico. Houve uma mudança sintática no português do Brasil e, dessa forma, a pesquisa mostra um progressivo enrijecimento da ordem em SN V³⁶. Mesmo com a diminuição das ocorrências de V SN no *corpus* sincrônico, encontramos a seguinte conclusão da pesquisadora:

[...] a quase categoricidade com que a ausência de concordância se associa a V SN e a manutenção desse resultado em relação aos demais fatores confirmam sua relevância e justificam a seleção do fator como explanatoriamente forte. Essa associação tem servido de argumento favorável à tese do caráter “menos subjetivo” do sujeito (ou SN) que ocorre em V SN, uma vez que a concordância verbal sempre foi considerada uma das propriedades formais definidoras da função-sujeito. (BERLINCK, 1988, p. 97-98).

É sabido que muitos casos de SN posposto ao verbo ocorrem com verbos apresentacionais. Castilho (2010) não concorda em classificar esse SN como sujeito nem como objeto da sentença. Baseado no conceito de caso gramatical (nominativo, acusativo, dativo, ablativo), Castilho (2010, p. 288) postula o absolutivo “como o caso abstrato projetado pelos verbos monoargumentais apresentacionais”.

Diante do exposto e com base em resultados de vários trabalhos (cf. LEMLE e NARO, 1977; NARO, 1981; RODRIGUES, 1987; BERLINCK, 1988; VIEIRA, 1995; MONGUILHOTT, 2001; GAMEIRO, 2005; MONTE, 2007; entre outros), nossa hipótese é a de que o sujeito/SN anteposto ao verbo tende a favorecer a concordância. O sujeito/SN posposto corresponde à categoria menos saliente, portanto acreditamos no baixo índice de concordância para essa configuração.

► Traço semântico do sujeito/SN

Esse grupo de fatores foi estabelecido por Scherre e Naro (1998a). Assim como os autores, estabelecemos a hipótese de que o sujeito/SN com o traço semântico [+ humano] favorece a presença da concordância verbal e o sujeito/SN com o traço semântico [- humano] favorece a ausência de concordância.

Além desses dois traços semânticos, também controlamos a animacidade (cf. VIEIRA, 1995, GAMEIRO, 2009; RUBIO, 2008, 2012). A expectativa é que os

³⁶ SN (sintagma nominal) V (verbo).

contextos com SN inanimado sejam desfavorecedores do emprego da marca de plural nos verbos.

As categorias consideradas para essa variável são as seguintes:

a) **SN [+ humano / + animado]**

(120) as pessoas **fogem** do campo... (M1PE)

(121) os locutores daqui **vão** pra Ribeirão Preto (F2PB)

b) **SN [- humano / + animado]**

(122) porque::: tenho pena dos cavalos... porque eles **vão** pra lá (F0PE)

(123) os passarinho **cantava**... (F0PB)

c) **SN [- humano / - animado]**

(124) ahn... não **é** os clubes... não conheço muitos clubes brasileiros...
(M2PE)

(125) até os brinquedo **era** diferente né? (M1PB)

Scherre e Naro (1998a) ampliam a hipótese da saliência para o nível semântico. O traço semântico [+ humano] é mais saliente do que o traço semântico [- humano]. Segundo Lucchesi, Baxter e Silva (2009, p. 354), “o favorecimento da concordância verbal quando o sujeito se refere a seres humanos está ligado à ideia de *agente*. Aquele que pratica a ação relaciona-se diretamente à ação, logo o verbo tenderia a concordar mais com ele”.

► **Distância entre o sujeito/SN e o verbo em termos do número de sílabas**

Naro e Scherre (1999a) decidiram medir a distância sujeito/SN-verbo em termos de sílabas, ao invés de palavras ou outra unidade hierarquicamente superior, porque cada sílaba ocupa aproximadamente a mesma quantidade de tempo, enquanto outras unidades têm dimensões diversas e diferenciadas.

Seguindo esses autores, o grupo de fatores focaliza o número de sílabas que separa os sujeitos/SNs antepostos de seus respectivos verbos. Vejamos as categorias:

a) de zero a duas sílabas entre sujeito/SN e verbo:

(126) os paranaense **tomaru** conta de São Carlos... (F1PB)

(127) os lavradores **não queriam** (M0PE)

(128) porque meus pais sempre **conversaram** muito comigo (M2PB)

b) presença de três a cinco sílabas entre sujeito/SN e verbo:

(129) acontecendo algumas coisa na igreja que **tava** desagradano (F1PB)

(130) eles... sabe?... meio que **apadrinhava** aquelas criança... (F1PB)

(131) porque há amigos que depois **falam** melhores que otros... (M1PE)

c) presença de seis ou mais sílabas entre sujeito/SN e verbo:

(132) as mulher antigamente não **fazia** isso (M0PB)

(133) os professores lá onde ela trabalha **fala** (F2PB)

(134) aqueles painéis que se aproveita pa pôr por cima do/ do telhado que...
absorve o sol etc. (F1PE)

De acordo com Naro (1981) e Naro e Scherre (1999a), nossa hipótese é a de que quanto mais óbvia ou mais saliente for a relação sujeito/SN-verbo, mais marca explícita de plural no verbo será encontrada. Ou seja, quanto mais perto estiver o sujeito/SN do seu respectivo verbo, esperamos encontrar mais concordância verbal.

► Presença ou ausência do ‘que’ relativo ou complementizador entre o SN controlador da concordância e o verbo

No artigo “A relação verbo/sujeito: o efeito máscara do *que* relativo”, Naro e Scherre (2003a), valendo-se da oração *Tem lugares **que já esgotou***, lançam a seguinte pergunta: “será que o pronome relativo *que* funciona no sentido de

relembrar ao falante qual é o sujeito da oração, aumentando a saliência da relação sujeito/verbo, e aumentando por isso as chances de marcação plural explícita do verbo em relação a outros casos em que há o mesmo número de sílabas entre sujeito e verbo? Ou, ao contrário, será que o pronome relativo *que* tem o efeito de mascarar a relação sujeito/verbo, provocando menos marcas explícitas de plural do que outras partículas de uma única sílaba?”

A fim de responder a essa questão com base nos dados de nossos *corpora*, resolvemos, da mesma maneira, controlar a presença/ausência do *que* relativo ou complementizador. O grupo compreende duas categorias:

a) ausência do ‘que’

(135) porque eles **compraru** um maquinário muito caro (M0PB)

(136) os irmãos não **queriam...** (F1PE)

(137) eles me **liberaram** o fundo de garantia (M1PB)

(138) agora os cursos técnicos nem sempre **abrem...** (F2PE)

b) presença do ‘que’

(139) trabalhei lá com pessoas **que era** Impecável memo... (M0PE)

(140) tem pessoas **que vai** lá (F1PB)

(141) dois casados **que é** da parte do meu pai... (M1PE)

(142) tinha uns **que também tinha** família boa... (M2PB)

Seguindo Naro e Scherre (2003a), a nossa hipótese é a de que esse elemento se diferencia dos demais elementos intervenientes que possam ocorrer entre o SN controlador da concordância e o verbo. Acreditamos que a presença do ‘que’ inibe as chances de o verbo apresentar marca explícita de plural.

► Tipo estrutural do sujeito/SN

A proposta do grupo de fatores tipo estrutural do sujeito/SN se baseia nos estudos de Rodrigues (1987), Naro e Scherre (2000), Scherre e Naro (2000),

Monguilhott (2009) e Rubio (2012). Controlamos nove formas de representação do sujeito/SN.

a) **Pronome vocês**

(143) cêis **tão** ficano doido... (F0PB)

(144) vocês só **pagam** a uma pessoa que faz duas funções... (F2PE)

b) **Pronome pessoal 'eles/elas'**

(145) eles **agrediram**... o autocarro do Benfica com sacos cheios de tinta azul... (M1PE)

(146) elas **terminaram** o terceiro (M1PB)

c) **SN pleno simples**

(147) **foi** os bandidos (M1PB)

(148) mas os meus pais **tinham** comércio em Valverde... (F1PE)

d) **SN pleno nu**

(149) pessoas que **enrola** elas (F1PB)

(150) também às vezes lá **morava** pessoas... (F0PE)

e) **SN pleno composto**

(151) daqui mais uns anos os ciganos e os tenderos e os pretos... **tomam** conta do/ da Europa... (M1PE)

(152) meu pai minha mãe sempre **pegaram** no nosso pé (F2PB)

f) **Numeral (como núcleo ou adjunto)**

(153) as duas **são** casada (F0PB)

(154) **existem** três clubes aqui... em Portugal (M2PE)

g) **Quantificador indefinido**

(155) uns **trabalhava** na construção civil (M0PE)

(156) muitos... **prefere** mulheres assim (F1PB)

h) **Pronome demonstrativo**

(157) os que **aprendeu** mais os mai novo né?... (M0PB)

(158) aqueles que **dizem** que não... (M2PE)

i) **Sujeito nulo**

(159) entretanto *eles* não **apareceram** e \emptyset nem **disseram** nada... (F1PE)

(160) *eles* **gostaram** dela... \emptyset **perguntaram** se ela te/teria alguém que mora em São Paulo (M2PB)

Há uma contradição nas hipóteses estabelecidas em alguns trabalhos. Por exemplo, espera-se que o sujeito nulo favoreça a concordância, pois se o sujeito não se encontra explícito na frase, a flexão não é redundante e tende a ser utilizada. Já o sujeito explícito, ‘eles/elas’ por exemplo, levaria ao uso de formas verbais não marcadas, pois a perda da informação causada pelo apagamento da desinência número-pessoal dos verbos é compensada pelo uso do pronome lexical. Outros trabalhos mantêm a hipótese em relação ao sujeito nulo, mas criam outra expectativa em relação ao sujeito pronominal ‘eles/elas’, pois dificilmente vai ocorrer à direita do verbo.

Trata-se de uma variável complexa, pois outras variáveis estão diretamente relacionadas. Zilles (2000), em seu trabalho sobre a posposição do sujeito ao verbo no português falado no Rio Grande do Sul, controla a forma de expressão do SN e conclui que os sujeitos de VS tendem a ser expressos por pronome indefinido ou SN pleno, registrando-se uma forte restrição ao uso de pronome reto com esta ordem.

Como podemos perceber, há uma forte relação entre a posição do sujeito/SN e o tipo estrutural do sujeito/SN. Em relação ao pronome pessoal ‘eles/elas’ e o pronome ‘vocês’, além da posição em relação ao verbo, o paralelismo formal pode

interferir de uma maneira muito atuante, já que a forma de plural explícita sempre estará presente nesses elementos.

► **Tipo de verbo: verbo ‘ser’ versus outros verbos**

Em alguns trabalhos com dados do PE (cf. MOTA e VIEIRA, 2008; GANDRA, 2009; VIEIRA, 2012; BRANDÃO e VIEIRA, 2012), as construções com o verbo ‘ser’ são apontadas como desfavorecedoras da concordância verbal. Já no trabalho de Varejão (2006), a hipótese de que o verbo ‘ser’ favoreceria o cancelamento da marca de plural não se confirmou.

Assim como relata Varejão (2006), em várias etapas do nosso trabalho – durante as entrevistas, na fase de transcrição e levantamento dos dados – parecia-nos que o verbo ‘ser’ apresentava comportamento diferente dos demais verbos, principalmente no PE. Por esse motivo, resolvemos controlar o verbo ‘ser’ separadamente nas duas amostras. Foram consideradas duas categorias:

a) **verbo ‘ser’**

(161) **era** aquelas mesona daqui lá grandona né?... (F0PB)

(162) não **é** todas que/ que vai não... (todas = novelas) (M0PB)

(163) Novas Oportunidades... mas há umas que **é** pós laboral e outras laborais... estas **é** laborais... mas também há aqui quem teja a tirá pós laboral... que já não **é**... remunerado... (M1PE)

(164) o ano passado... as obras obrigatórias **foi** A Relíquia de Eça de Queirós... e::: ... A Relíquia... uhn... ah... o Sermão de Santo Antonio aos peixe do Padre António Vieira... ... e::: ... não sei... já não ah e o Frei do Luís de Sousa... Frei Luís de Sousa... (F2PE)

b) **outros verbos**

(165) foi a casa praticamente onde eles **nasceram**... (F1PE)

(166) então elas **tinham** oportunidade de estudá... (F1PB)

(167) os de cá de Évora **tão** na de honra acho eu... (referente: jogadores de rugby) (M2PE)

(168) alguns professores até **ficavam** meio chateados... (M2PB)

O 'ser' é um verbo complexo que normalmente ocorre em várias construções da língua portuguesa. Encontramos o verbo 'ser' em sentenças atributivas, equativas e apresentacionais. É importante lembrar que incluímos na nossa análise o verbo 'ser' em estruturas clivadas, pseudoclivadas, em orações com valor existencial e nas designações de tempo, distância, quantidade, valor. Sendo assim, a hipótese que norteia esse grupo de fatores é a de que o verbo 'ser' favorece a ausência de concordância verbal.

3.2.5.2 Grupos de fatores sociais (ou variáveis sociais)

GRUPOS DE FATORES	FATORES
1) Gênero	– masculino – feminino
2) Escolaridade	– não alfabetizados – ensino fundamental / ensino básico (EJA) – ensino médio / ensino secundário

► GÊNERO

A motivação para avaliar o papel do gênero sobre a variação linguística pode ser sintetizada pelas palavras de Rodrigues (1987):

As variedades linguísticas ligadas a sexo aparecem como resultado do fato de a língua, enquanto fenômeno social, estar intimamente relacionada com atitudes sociais. Homens e mulheres são socialmente diferentes, a eles são atribuídos diferentes papéis e, por isso, deles se esperam padrões de comportamento diferenciados [...] (RODRIGUES, 1987, p.199).

Segundo Paiva (2004), a primeira referência à correlação entre variação linguística e o fator gênero/sexo se encontra em Fischer (1974 [1958]) em um estudo intitulado *Social Influences on the Choice of a Linguistic Variant*³⁷.

Labov (2001) retoma o papel importante das mulheres nos processos de variação e mudança linguística e discute o ‘paradoxo do gênero’. Conforme o tipo de mudança, as mulheres apresentam um comportamento diferente. De acordo com Paiva (2004, p. 36), no estudo da correlação entre gênero/sexo e mudança linguística, um aspecto a considerar é o valor social da variante inovadora.

Quando se trata de implementar na língua uma forma socialmente prestigiada [...], as mulheres tendem a assumir a liderança da mudança. Ao contrário, quando se trata de implementar uma forma socialmente desprestigiada, as mulheres assumem uma atitude conservadora e os homens tomam a liderança do processo. (PAIVA, 2004, p. 36).

³⁷ Influências sociais na escolha de variantes linguísticas.

Em relação aos fenômenos de variação estável, estudos sociolinguísticos normalmente verificam que as mulheres tendem a se aproximar mais da variedade padrão do que os homens.

Por que os homens e as mulheres desempenham um papel distinto nos fenômenos de variação e mudança linguística? Para Chambers e Trudgill (1994 [1980]), não há uma explicação única, ou uma que seja amplamente aceita. Os autores vão apontar alguns fatores, como:

[...] las mujeres siguen teniendo menos oportunidades para triunfar, y marcan, por tanto, su estatus social por su apariencia y comportamiento (también lingüístico) más que por lo que hacen. [...] El mayor papel que las mujeres han tenido tradicionalmente en la socialización de los niños las lleva a ser más sensibles a las normas del comportamiento “aceptado”. [...] (CHAMBERS e TRUDGILL, 1994 [1980], p. 133-134).

Levando em consideração que a variação da concordância verbal de 3ª pessoa do plural envolve uma variante socialmente prestigiada e outra desprestigiada, a nossa hipótese é a de que as mulheres vão apresentar mais a marca explícita de plural nos verbos, ou seja, vão usar mais a variante de prestígio do que os homens.

► ESCOLARIDADE

Parece comumente aceita entre os linguistas brasileiros a ideia de que o grau de escolaridade é um bom indicador para ajudar a esclarecer as diferenças entre o português padrão e o português popular. Segundo Oliveira e Silva & Paiva (1996, p.343), “em países em que a escolarização não é apanágio das(s) classe(s) mais favorecidas(s), é a variável classe social que tem preferentemente servido de marcador”³⁸.

Para Rodrigues (1987), o baixo nível de escolaridade é decisivo para a identificação dos usuários de uma variedade popular de língua falada. Mas a pesquisadora salienta que, ao postular a existência de um segmento populacional urbano caracterizado por um baixo nível de escolaridade, não está omitindo outros atributos sociais que caracterizam esse grupo sociolinguístico popular.

³⁸ As autoras citam os estudos de Labov (1962, 1966), de Wolfram (1969) e de Trudgill (1974).

A mesma ideia também permeia o trabalho de Bortoni-Ricardo (2008 [1981], p. 374-375) quando afirma ter usado, como indicador de classe social, o nível escolar, pois ele reflete outras variáveis de ordem social e econômica. De acordo com a pesquisadora, temos “falta de estudos sociológicos que apresentem uma estratificação da população da cidade – problema, aliás, com que se defrontam sempre as pesquisas sociolinguísticas no Brasil”.

Oliveira e Silva & Paiva (1996), apresentando os condicionamentos extralinguísticos segundo resultados de diversos estudos, concluem:

De tudo o que vimos, as evidências apontam, por conseguinte, para um papel não trivial da escola pública na compleição linguística dos indivíduos. Com efeito, seja direta seja indiretamente, a participação da escola acaba sendo decisiva na modificação do comportamento linguístico.

Portanto, em vez de minimizar o efeito da escolarização no uso da língua, cabe analisar criticamente a interferência decisiva da escola na configuração linguística da comunidade. (OLIVEIRA e SILVA & PAIVA, 1996, p. 350).

Cabe lembrar que em Portugal, como bem ressalta Maria Helena Mira Mateus (2006), a distância entre o que se poderia chamar o português popular (ainda que não se use essa denominação) e as variedades de prestígio não é tão grande como no Brasil.

Como expusemos na subseção 3.2.1, o nosso estudo contempla três níveis diferenciados de escolaridade, a saber: (i) informantes não alfabetizados, (ii) concluintes do ensino fundamental/básico na EJA (Educação de Jovens e Adultos) e (iii) concluintes do ensino médio/secundário.

A nossa hipótese é a de que, quanto maior for o grau de escolaridade, maiores são as chances de o falante estar no grupo dos que falam o português padrão, mesmo a escolaridade sendo supletiva. Não se pode deixar de salientar que um dos grupos dos escolarizados dos nossos *corpora* é formado por pessoas com formação escolar supletiva (EJA – Educação de Jovens e Adultos).

4 ANÁLISE DOS DADOS: APRESENTANDO E DISCUTINDO OS RESULTADOS

4.1 Notas introdutórias

Nesta seção, apresentamos os resultados obtidos na análise dos dados da amostra da cidade de São Carlos (PB) e da cidade de Évora (PE).

Feitas todas as transcrições das entrevistas, os dados foram levantados e, após a codificação conforme as variáveis linguísticas e sociais estabelecidas, submetidos ao programa estatístico Goldvarb-X, que realiza uma análise multivariada (cf. TAGLIAMONTE, 2006, 2007). O Goldvarb-X é uma das versões para ambiente *Windows* do pacote de programas computacionais denominado *Varbrul* (cf. PINTZUK, 1988; SANKOFF, 1988;).

Guy e Zilles (2007) explicam o que é uma análise multivariada:

A análise se chama 'multivariada' porque permite investigar situações em que a variável linguística em estudo é influenciada por vários elementos do contexto, ou seja, múltiplas variáveis independentes. A investigação mede os efeitos, bem como a significância dos efeitos, dessas variáveis independentes sobre a ocorrência das realizações da variável que está sendo tratada como dependente. (GUY e ZILLES, 2007, p. 105).

Dessa forma, o programa estatístico Goldvarb-X nos forneceu as seguintes informações:

- (1) a frequência geral de aplicação e de não aplicação da regra de concordância verbal de terceira pessoa do plural, bem como as frequências associadas às variáveis estabelecidas³⁹;
- (2) a seleção por ordem de relevância das variáveis com significância estatística (rodada *step-up*);
- (3) as variáveis descartadas como irrelevantes, ou seja, sem significância estatística (rodada *step-down*);
- (4) o cruzamento das variáveis (função *Cross Tabulation*);
- (5) os pesos relativos, que permitem uma visão mais apurada dos dados.

³⁹ Tomamos por base a aplicação da regra de concordância verbal. Sendo assim, os resultados estatísticos que apresentamos devem ser lidos nesse sentido.

Os pesos relativos são valores que vão de zero a um e que indicam matematicamente o peso com que um fator (linguístico ou extralinguístico) influencia o uso de uma variante, em relação a todos os fatores levados em conta na observação de um fenômeno de variação linguística. Quando o peso relativo de um fator é próximo de zero, significa que tal fator desfavorece o uso da variante. Quando o peso relativo é igual a 0,50, significa que ele não está correlacionado ao uso da variante – tal valor é, pois, o ponto denominado neutro. Finalmente, quanto mais próximo for de 1 (um), maior será o peso com que o fator favorece o uso da variante. Os valores dos pesos relativos são obtidos a partir de fórmulas estatísticas complexas, do tipo de regressão logística, em que se comparam as porcentagens com que os dados se distribuem pelos diferentes fatores. (BELINE, 2005, p. 132).

Além da ordem de seleção das variáveis (ou grupos de fatores), outra evidência que podemos utilizar para interpretar a força de cada variável é o *range*. Segundo Tagliamonte (2006, p. 242)⁴⁰, a força de cada grupo de fatores é medida pelo *range*, que é então comparado com os *ranges* de outros grupos de fatores significativos.

O *range* é calculado pela diferença entre o peso relativo mais alto e o peso relativo mais baixo. Quando estes números são comparados para cada grupo de fatores em uma análise, o valor maior (o *range*) identifica a restrição mais forte. [...] O *range* (ou a magnitude do efeito) nos permite situar um grupo de fatores em relação a outro. (TAGLIAMONTE, 2006, p. 242, tradução Naro e Scherre, 2010)⁴¹.

Ainda nos termos de Tagliamonte (2006, p. 251), “o valor do *range* não é um peso relativo. Ele é apenas um número e não deve aparecer com decimal”⁴².

Iniciamos apresentando os resultados com os dados do PE e depois os resultados com os dados do PB. A ênfase da nossa análise recairá nos grupos de fatores selecionados pelo programa Goldvarb-X, mas não deixaremos de apresentar os resultados dos grupos de fatores que não foram selecionados. Na última parte desta seção, retomamos as variáveis mais significativas nas amostras do PE e do PB, tecendo uma comparação entre as duas variedades estudadas.

⁴⁰ Além do livro de Tagliamonte (2006), consultamos também Naro e Scherre (2010).

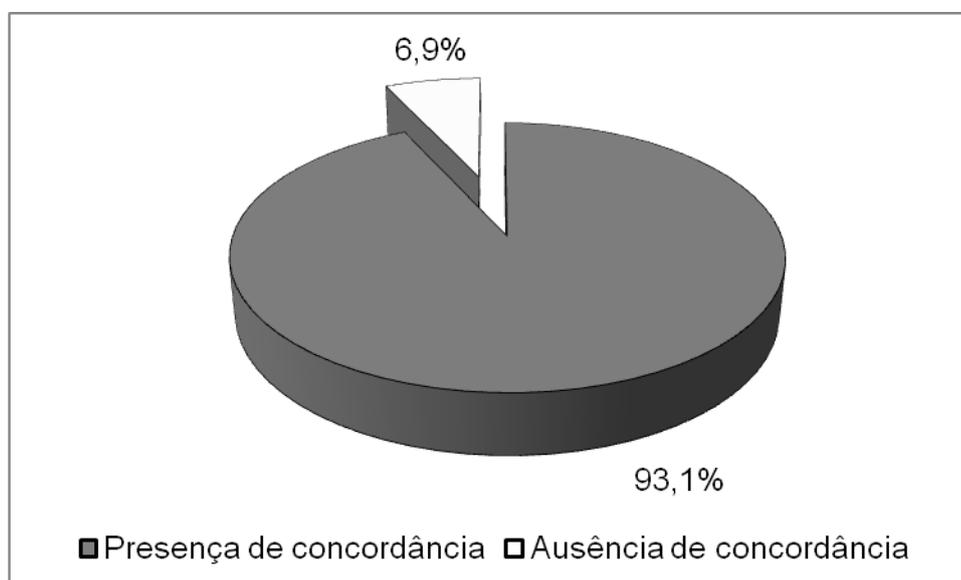
⁴¹ The range is calculated by subtracting the lowest factor weight from the highest factor weight. When these numbers are compared for each of the factor groups in an analysis, the highest number (i.e. range) identifies the strongest constraint. [...] The range (or magnitude of effect) enables you to situate factor groups with respect to each other. (TAGLIAMONTE, 2006, p. 242).

⁴² The range value is not a factor weight. It is simply a number and should not appear with a decimal. (TAGLIAMONTE, 2006, p. 251).

4.2 Análise e discussão dos resultados gerais do PE

Em Évora (PE), encontramos um índice muito alto de frequência de concordância verbal de terceira pessoa do plural na fala informal das pessoas. Do total de 1.440 ocorrências, apenas 100 (6,9%) trazem a variante zero de plural nos verbos, sendo que 1.340 (93,1%) apresentam a marca explícita de plural nos verbos, ou seja, a concordância verbal.

Gráfico 1: Distribuição geral dos dados do PE.



A presença da concordância verbal apresentou *input* 0,964 e a rodada significância 0,015. Os grupos de fatores selecionados pelo programa Goldvarb-X, na ordem de relevância, foram: posição do sujeito/SN em relação ao verbo, traço semântico do sujeito/SN, tipo estrutural do sujeito/SN, tipo de verbo (verbo 'ser' *versus* outros verbos), saliência fônica e gênero.

Realizamos outra rodada com o programa Goldvarb-X sem as ocorrências com o verbo 'ser', que será discutida no final desta seção.

4.2.1 Posição do sujeito/SN em relação ao verbo⁴³

A posição do sujeito/SN em relação ao verbo foi o primeiro grupo de fatores selecionado pelo programa Goldvarb-X.

Tabela 3: Frequência e peso relativo de concordância verbal segundo a posição do sujeito/SN em relação ao verbo no PE.

Fatores	Frequência	PR
– sujeito/SN anteposto	680/715 = 95,1%	0,581
– sujeito/SN posposto	154/199 = 77,4%	0,237
Total	834/914 = 91,2%	Range 344
<i>Input: 0,964</i>		
<i>Significância: 0,015</i>		

O índice de frequência de concordância diminui de 95,1% para 77,4% quando o sujeito/SN está posposto ao verbo. É uma diferença de 17,7 pontos percentuais. A diferença entre as duas categorias chega a 344 em termos de peso relativo.

Os exemplos (1-6) ilustram a ausência de concordância verbal na amostra do PE com a posposição do sujeito/SN:

- (1) não me lembro o nome delas... .. **é os janeiros...** ((risos)) (FOPE)
- (2) dantes **trabalhava** lá **dez ou doze pessoas...** (M0PE)
- (3) morava em Estremoz... **vinha os comboios...** cheinhos de pessoas pa virem ao São João... (F1PE)
- (4) aí **nasce as cooperativas...** (M1PE)
- (5) essencialmente **é estes três...** o Porto... Sporting e o Benfica... (M2PE)
- (6) **existe sites** onde dá pa vê filmes... (M2PE)

Outros trabalhos também revelam a importância da posição do sujeito/SN em relação ao verbo como fator condicionante da variação da concordância verbal de terceira pessoa do plural no PE (cf. VAREJÃO, 2006; NARO e SCHERRE, 2007; MONGUILHOTT, 2009; BAZENGA, 2011; RUBIO, 2012; VIEIRA, 2012; BRANDÃO e VIEIRA, 2012).

⁴³ No controle desta variável, foram desconsideradas as 526 ocorrências com sujeito nulo.

Para efeito de comparação, vejamos os resultados encontrados por Monguilhott (2009).

Tabela 4: Frequência e peso relativo de concordância verbal em função da posição do sujeito/SN em relação ao verbo no PE, no *corpus* sob análise e no *corpus* de Monguilhott (2009).

Fatores	Frequência de concordância			
	Nossos resultados		Monguilhott (2009)	
	Frequência	PR	Frequência	PR
– sujeito/SN anteposto	680/715 = 95,1%	0,581	354/381 = 92%	0,56
– sujeito/SN posposto	154/199 = 77,4%	0,237	43/72 = 59%	0,18
Total	834/914 = 91,2%		397/453 = 87%	

Os nossos resultados em termos de peso relativo são muito semelhantes aos encontrados pela pesquisadora, confirmando-se a relevância da posposição do sujeito/SN para a explicação da variação.

No ano letivo 2009-2010, o Ministério da Educação de Portugal deu continuidade ao projeto dos testes intermédios. Em janeiro de 2010, mais de 1200 escolas com 3º ciclo do ensino básico realizaram o teste intermédio de Língua Portuguesa. O teste aplicado aos estudantes do 9º ano apresentou dois grupos de itens. No Grupo I, avaliaram-se competências nos domínios da leitura e da escrita. No grupo II, avaliaram-se competências do domínio do funcionamento da língua. Nosso interesse é mostrar o resultado do item 4.1 do Grupo II. Os estudantes tinham de indicar a função sintática desempenhada pela expressão sublinhada na seguinte frase: “Nessa manhã, chegaram os passageiros ao navio”. Apenas 35% dos estudantes responderam corretamente, de acordo com a tradição gramatical. A maioria (65%) não reconheceu “os passageiros” como o sujeito da frase. Como a questão foi aberta e não de múltipla escolha, não disponibilizaram as outras respostas que os estudantes deram. A dificuldade manifestada na questão justifica-se pelo fato de o sujeito aparecer posposto ao verbo.

4.2.2 Traço semântico do sujeito/SN

No nosso estudo com dados do PE, o grupo de fatores traço semântico do sujeito/SN foi o segundo selecionado pelo programa Goldvarb-X. Como hipotetizado,

vemos que há maiores chances de aplicação da concordância com sujeito [+ humano] e [+ animado].

Tabela 5: Frequência e peso relativo de concordância verbal em função do traço semântico do sujeito/SN no PE.

Fatores	Frequência	PR
– SN [+ humano / + animado]	1.082/1.130 = 95,8%	0,555
– SN [– humano / + animado]	28/31 = 90,3%	0,438
– SN [– humano / – animado]	230/279 = 82,4%	0,295
Total	1.340/1.440 = 93,1%	Range 260
<i>Input: 0,964</i>		
Significância: 0,015		

Quando temos o sujeito/SN com o traço semântico [+ humano / + animado], como em (7-8), o índice de frequência de concordância é de 95,8%, com 0,555 de peso relativo.

- (7) há deputados que **contam** anedotas... (F2PE)
 (8) as mulheres hoje **vão** pa tropa (M0PE)

Na categoria intermediária, representada por sujeito/SN [– humano / + animado], conforme exemplos (9-10), encontramos a frequência (90,3%) e o peso relativo (0,438) menores do que com sujeito/SN com traço [+ humano], mas esses índices ainda são maiores quando comparamos com a categoria de sujeito/SN inanimado.

- (9) e então tínhamos... tínhamos os/ os gatos e os cães e as galinhas e os perus... e **era** os nossos companheiros... (F0PE)
 (10) é onde esses porcos **são** engordados... (M1PE)

Com sujeito/SN inanimado, como nas ocorrências (11-12), as chances de aplicação da regra de concordância são menores (82,4% de frequência e 0,295 de peso relativo).

- (11) já tem se feito (algumas) greves mas nunca **deu...** nunca **deu** resultado... (M2PE)
- (12) também migas também **é** muito boas... (F0PE)

Naro e Scherre (2007), Monguilhott (2009) e Bazenga (2011) analisaram apenas a distinção [+ humano] e [- humano]. Os resultados de suas pesquisas também revelam a influência dessa variável semântica. No trabalho de Monguilhott (2009), esse grupo de fatores foi selecionado como estatisticamente relevante em primeiro lugar.

Controlamos esse grupo de fatores da mesma forma que foi controlado por Rubio (2012). Na tabela 6, a seguir, apresentamos os resultados obtidos com base nos dados de nossa amostra, comparando-os aos resultados obtidos por Rubio (2012).

Tabela 6: Frequência e peso relativo de concordância verbal segundo o traço semântico do sujeito/SN no PE, na amostra sob análise e na amostra de Rubio (2012).

Fatores	Frequência de concordância			
	Nossos resultados		Rubio (2012)	
	Frequência	PR	Frequência	PR
- SN [+ humano / + animado]	1.082/1.130 = 95,8%	0,555	733/756 = 97%	0,588
- SN [- humano / + animado]	28/31 = 90,3%	0,438	24/26 = 92,3%	0,326
- SN [- humano / - animado]	230/279 = 82,4%	0,295	282/325 = 86,8%	0,316

Rubio (2012) não encontra diferença significativa entre a categoria com sujeito/SN [- humano / + animado] e a categoria com SN [- humano / - animado]. Os contextos com sujeito/SN [+ humano/ + animado] se destacam dos demais, favorecendo a marca de plural nos verbos.

Calculando a diferença entre o peso relativo mais alto e o peso relativo mais baixo, temos um *range* de 272 com os dados de Rubio (2012) e um *range* de 260 com os nossos dados. Os *ranges* nas duas pesquisas são muito semelhantes.

Realizamos o cruzamento da variável traço semântico do sujeito/SN com a variável posição do sujeito/SN em relação ao verbo, que foi a primeira selecionada pelo Goldvarb-X. A tabela 7 traz os resultados obtidos:

Tabela 7: Frequência de concordância verbal segundo o traço semântico do sujeito/SN e a posição do sujeito/SN em relação ao verbo no PE.

Fatores	Frequência de concordância	
	– sujeito/SN anteposto	– sujeito/SN posposto
– SN [+ humano / + animado]	541/558 = 97%	78/96 = 81%
– SN [– humano / + animado]	15/16 = 94%	3/3 = 100%
– SN [– humano / – animado]	124/141 = 88%	73/100 = 73%

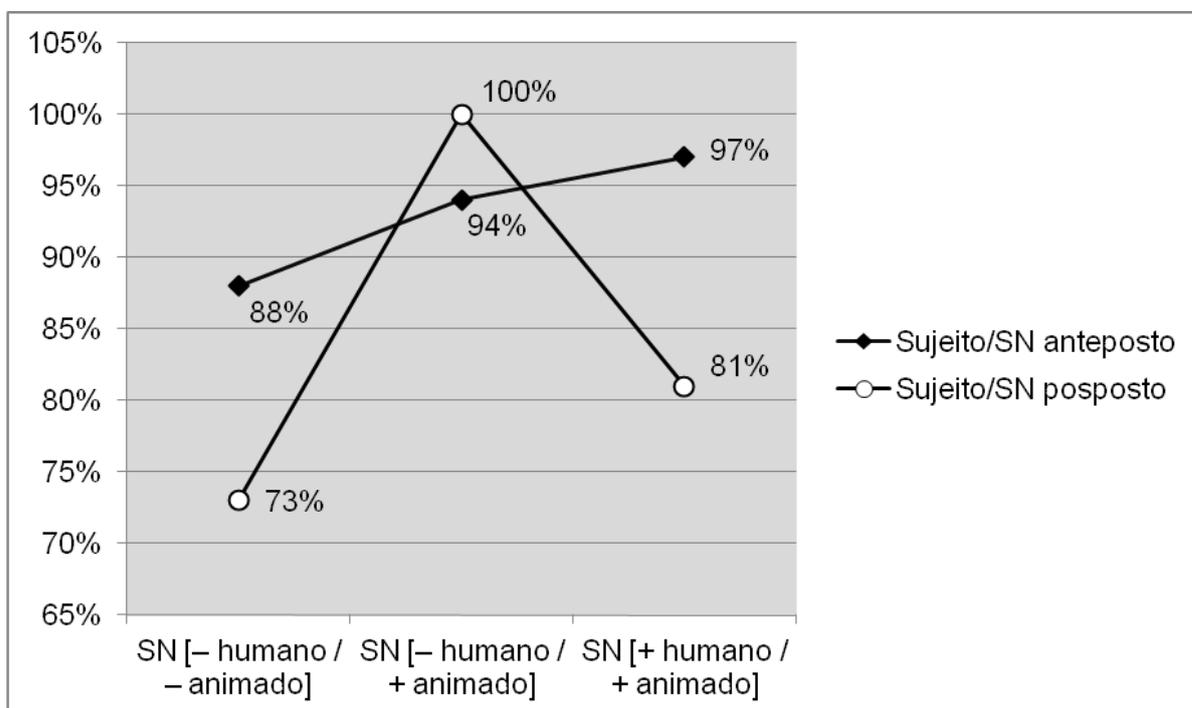
Com o sujeito/SN [+ humano / + animado] e anteposto ao verbo, encontramos um índice de frequência de concordância maior (97%) do que quando o sujeito/SN [+ humano / + animado] vem posposto ao verbo (81%). Observamos também que os casos com o sujeito/SN inanimado anteposto favorecem mais a marca explícita de plural no verbo (88%) do que quando o sujeito/SN inanimado está posposto (73%).

A categoria com sujeito/SN [– humano / + animado] apresenta um comportamento diferente: 94% de frequência de concordância com o sujeito/SN anteposto e 100% de frequência com o sujeito/SN posposto. No entanto, temos de considerar que há poucas ocorrências nessa categoria. São 16 dados com o sujeito/SN anteposto e apenas 3 dados com o sujeito/SN posposto. Esses 3 dados com o sujeito/SN posposto trazem a marca explícita de plural no verbo:

- (13) **eram** vacas alentejanas vermelhas... (M0PE)
- (14) **são** aqueles que vivem no montado... à vontade... (aqueles = porcos alentejanos) (M1PE)
- (15) onde é que **são**... criados esses animais... (M1PE)

O próximo gráfico nos permite uma visualização melhor da atuação desses dois grupos de fatores.

Gráfico 2: Frequência de concordância verbal segundo o traço semântico do sujeito/SN e a posição do sujeito/SN em relação ao verbo no PE.



Em termos de frequência, os resultados expressos na tabela 7 e no gráfico 2 revelam pouca interação e mostram a relevância desses dois grupos de fatores na pouca variação existente na concordância verbal de 3ª pessoa do plural no PE.

4.2.3 Tipo estrutural do sujeito/SN

O terceiro grupo de fatores selecionado pelo programa Goldvarb-X foi o tipo estrutural do sujeito/SN. Os resultados vêm expressos na tabela 8, a seguir:

Tabela 8: Frequência e peso relativo de concordância verbal de acordo com o tipo estrutural do sujeito/SN no PE.

Fatores	Frequência	PR
– Pronome <i>vocês</i>	8/8 = 100%	–
– Pronome pessoal <i>eles/elas</i>	168/169 = 99,4%	0,805
– Numeral (núcleo ou adjunto)	76/85 = 89,4%	0,664
– Sujeito nulo	506/526 = 96,2%	0,527
– Pronome demonstrativo	35/37 = 94,6%	0,462
– Quantificador indefinido	72/78 = 92,3%	0,436
– SN pleno simples	390/436 = 89,4%	0,370
– SN pleno nu	71/84 = 84,5%	0,294
– SN pleno composto	14/17 = 82,4%	0,197
Total	1.340/1.440 = 93,1%	Range 608
<i>Input: 0,964</i>		
<i>Significância: 0,015</i>		

A categoria que apresenta a menor frequência (82,4%) e o menor peso relativo (0,197) é a categoria de SN pleno composto. Essa categoria estava, inicialmente, dividida em duas: SN composto com o núcleo mais próximo do verbo no singular e SN composto com o núcleo mais próximo do verbo no plural. Com os poucos dados em cada uma, resolvemos amalgamá-las. Ainda assim, ficamos com uma categoria com poucas ocorrências. São 17 dados, sendo que três não trazem a marca de plural no verbo:

- (16) os miúdos (o::s)... graúdos... miúdos... **junta-se** pa fazê grandes festas no um de dezembro... (F1PE)
- (17) **é** ciganos... tenderos... (coisas) assim... (M2PE)
- (18) se formos pa ciências e tecnologias **é** sempre químicas e física(s)... essas coisas assim... (F2PE)

Como podemos observar, das três ocorrências sem a marca explícita de plural no verbo, duas (17-18) são com o SN posposto e com o verbo 'ser'

apresentacional, fatores que favorecem no PE a não concordância. Vale destacar que esses dois dados de SN composto posposto são os únicos na nossa amostra do PE.

Desfavorecendo as marcas de concordância verbal, encontramos também o SN pleno simples, com 89,4% de frequência e 0,370 de peso relativo, e o SN pleno nu, com 84,5% de frequência e 0,294 de peso relativo. A hipótese se confirma, pois essas estruturas são mais propensas de apresentar a posposição do sujeito/SN. Esse fato foi evidenciado por Zilles (2000, p.83)⁴⁴: “o sujeito de VS tende a ser expresso por SN pleno ou pronome indefinido”. Em relação aos quantificadores indefinidos, podemos observar que também tendem a desfavorecer a concordância verbal (92,3% de frequência e 0,436 de peso relativo). O comportamento dos pronomes demonstrativos (94,6% de frequência e 0,462 de peso relativo) é semelhante ao dos quantificadores indefinidos.

Favorecendo a marcação da concordância verbal, temos o pronome ‘você’, o pronome pessoal ‘eles/elas’, o numeral e o sujeito nulo. As oito ocorrências com o pronome ‘você’ trazem a marca de plural no verbo. Quando o sujeito é representado pelo pronome eles/elas, a regra se apresenta praticamente categórica no emprego da concordância verbal, com frequência de 99,4% e peso relativo de 0,805. Há apenas um dado sem a marca de plural no verbo:

(19) eles **abala** os dois no carro... (F0PE)

Esse dado é de uma senhora analfabeta e, como podemos notar, ela usa o pronome eles anteposto ao verbo e reforça o sujeito com um quantificador definido (os dois) depois do verbo.

Quando há um sintagma incluindo um numeral, como núcleo ou adjunto, os nossos dados revelam um contexto favorecedor do emprego da concordância verbal. É interessante notar que a frequência (89,4%) não é tão alta, comparando com outras categorias que apresentam valores de frequência maiores. No entanto, o peso relativo é forte: 0,664. Esses resultados mostram a importância de se separar os quantificadores indefinidos dos quantificadores definidos nos estudos desta natureza.

⁴⁴ VS (verbo-sujeito).

Por fim, verificamos que o sujeito nulo tende a favorecer a presença de concordância (96,2% de frequência e 0,527 de peso relativo), pois, com a ausência do sujeito, a forma verbal marcada em número e pessoa é uma pista formal e semântica para identificá-lo.

Esse grupo de fatores também foi selecionado como estatisticamente relevante na pesquisa de Rubio (2012) e vários fatores exibem comportamento semelhante ao observado por ele. Destacamos em seu trabalho o peso relativo de 0,807 para o fator pronome pessoal, 0,566 para o sujeito nulo, 0,348 para o SN pleno simples e 0,186 para o SN pleno nu. No trabalho de Monguilhott (2009), o programa Goldvarb-X não selecionou o tipo estrutural do sujeito, mas, pelas frequências puras, vemos semelhanças significativas. Em sua pesquisa, os sujeitos pleno simples, pleno nu e pleno composto foram os que menos preservaram as marcas de concordância. Já o sujeito pronome pessoal e o sujeito nulo tiveram o maior percentual de marcação da concordância verbal.

É importante lembrar que, dentre os poucos trabalhos variacionistas sobre o tema no PE que utilizam o programa Goldvarb-X na análise, este grupo de fatores obteve significância estatística somente no de Rubio (2012) e no nosso.

4.2.4 Tipo de verbo: verbo 'ser' versus outros verbos

Os valores expostos na próxima tabela evidenciam que o grupo de fatores tipo de verbo (verbo 'ser' versus outros verbos) é relevante nos estudos de variação da concordância verbal no PE. Na ordem de relevância, este grupo de fatores foi o quarto selecionado.

Tabela 9: Frequência e peso relativo de concordância verbal segundo o tipo de verbo (verbo 'ser' versus outros verbos) no PE.

Fatores	Frequência	PR
– verbo 'ser'	291/342 = 85,1%	0,247
– outros verbos	1.049/1.098 = 95,5%	0,586
Total	1.340/1.440 = 93,1%	Range 339
<i>Input: 0,964</i>		
<i>Significância: 0,015</i>		

Sentenças com o verbo 'ser' inibem, realmente, a marca explícita de plural nos verbos. A diferença entre as duas categorias é de 10,4 pontos percentuais em termos de frequência e 339 em termos de peso relativo. Lembrando que alguns trabalhos (cf. MOTA e VIEIRA, 2008; GANDRA, 2009; VIEIRA, 2012; BRANDÃO e VIEIRA, 2012) já apontam um comportamento diferenciado em estruturas com o verbo 'ser'.

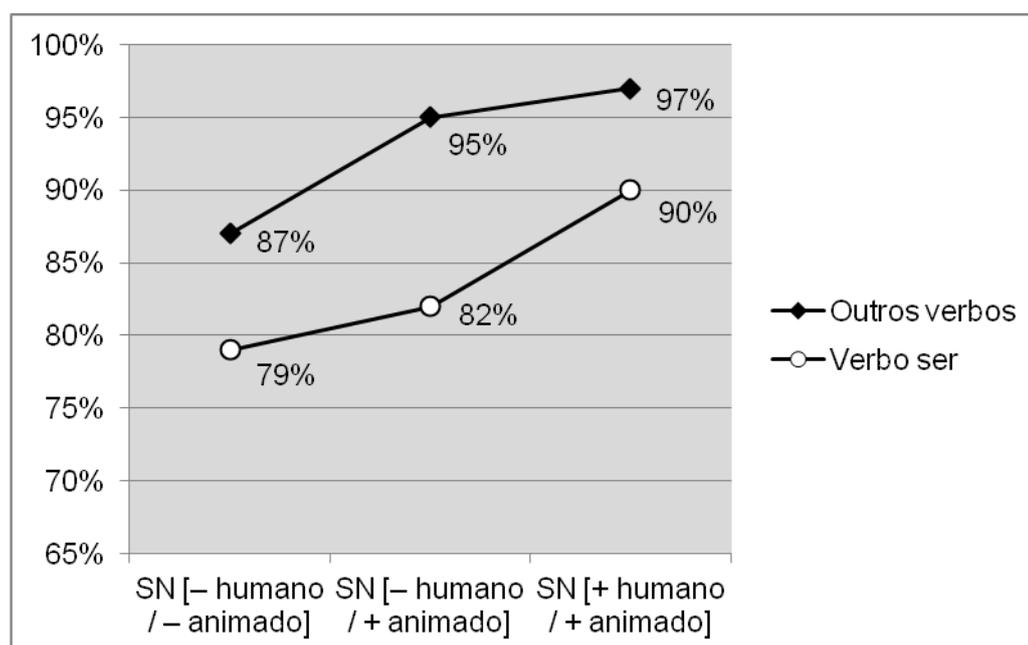
Com o objetivo de compreender melhor o comportamento do verbo 'ser', realizamos o cruzamento com os grupos de fatores traço semântico do sujeito/SN e posição do sujeito/SN em relação ao verbo.

Primeiramente, apresentamos o cruzamento do tipo de verbo (verbo 'ser' *versus* outros verbos) com o traço semântico do sujeito/SN. A tabela e o gráfico a seguir permitem observar o comportamento dessas variáveis.

Tabela 10: Frequência de concordância verbal segundo o tipo de verbo (verbo 'ser' *versus* outros verbos) e o traço semântico do sujeito/SN no PE.

Fatores	Frequência de concordância	
	– outros verbos	– verbo 'ser'
– SN [+ humano / + animado]	914/944 = 97%	168/186 = 90%
– SN [– humano / + animado]	19/20 = 95%	9/11 = 82%
– SN [– humano / – animado]	116/134 = 87%	114/145 = 79%

Gráfico 3: Frequência de concordância verbal segundo o tipo de verbo (verbo 'ser' *versus* outros verbos) e o traço semântico do sujeito/SN no PE.



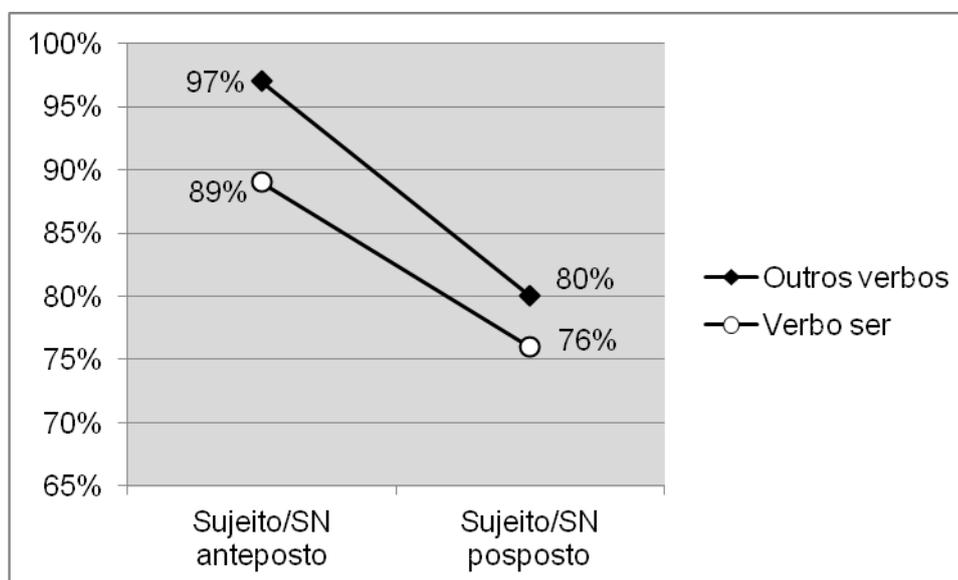
Conforme os resultados acima, verificamos o favorecimento ao emprego da marca de plural com sujeito/SN [+ humano]. O índice de frequência de concordância cai um pouco com sujeito/SN [- humano / + animado], e cai um pouco mais quando temos o sujeito/SN inanimado. O que chama a atenção é que essa tendência é menor nas três categorias da variável semântica quando o verbo 'ser' é empregado.

A próxima tabela e o próximo gráfico retratam os resultados do cruzamento das variáveis tipo de verbo (verbo 'ser' *versus* outros verbos) e posição do sujeito/SN em relação ao verbo.

Tabela 11: Frequência de concordância verbal segundo o tipo de verbo (verbo 'ser' *versus* outros verbos) e a posição do sujeito/SN em relação ao verbo no PE.

Fatores	Frequência de concordância	
	– sujeito/SN anteposto	– sujeito/SN posposto
– verbo 'ser'	142/160 = 89%	87/115 = 76%
– outros verbos	538/555 = 97%	67/84 = 80%

Gráfico 4: Frequência de concordância verbal segundo o tipo de verbo (verbo 'ser' *versus* outros verbos) e a posição do sujeito/SN em relação ao verbo no PE.



Sem interação, vemos o efeito das duas variáveis atuando na variação da concordância verbal no PE. Observamos que a frequência de concordância diminui, tanto com o verbo 'ser' quanto com outros verbos, quando o sujeito/SN está posposto ao verbo. No entanto, o índice ainda é menor com o verbo 'ser' (76%).

Salientamos que o número de ocorrências de sujeito/SN posposto com o verbo 'ser' é maior (115) comparado ao número de ocorrências de sujeito/SN posposto com outros verbos (84). Verificando os dados, percebemos que a grande maioria é com o verbo 'ser' em construções apresentativas, como ilustram os exemplos abaixo:

- (20) no fim já **era** tratores... (F0PE)
- (21) trabalhá na máquinas... geralmente **era** os homens (M0PE)
- (22) **é** as tendinites... (F1PE)
- (23) éh... **é** jogos... porreros que gosto... da internet... (M1PE)
- (24) ma não é/ não **é** os ciganos... (F2PE)
- (25) **era** rapazes do::s oito aos dezessete... ciganos... todos numa/ numa sala... (M2PE)

Como já nos referimos anteriormente, Castilho (2010, p. 288) sugere o termo absolutivo para esses SNs introduzidos pelos verbos apresentacionais. Segundo o autor, "é graças a essa categoria gramatical que podemos representar em nossa língua os tópicos discursivos novos, que precisamos incluir na corrente da fala". Para Castilho (2010) e Bagno (2011), o SN que é introduzido por um verbo apresentacional não tem propriedades sintáticas de sujeito nem de objeto.

Chama a atenção, também, o alto índice de frequência de concordância (97%) na categoria outros verbos com o sujeito/SN anteposto. É a situação com o maior número de ocorrências (555) e apenas 17 dados não trazem a marca explícita de plural no verbo.

Cabe ainda ressaltar a variação da concordância do verbo 'ser' no fenômeno da clivagem. Com o foco na 3ª pessoa do plural, foram encontradas na amostra estudada 22 ocorrências (19 clivadas e 3 pseudoclivadas), sendo que cinco clivadas não trazem o verbo 'ser' com a marca de plural.

- (26) **era** uma(s) ceiferas... que **cefavam** só (M0PE)
- (27) **era** as pessoas que **tinham** lá famílias a trabalhá já... (M0PE)
- (28) já não **é** as pessoas que **mandam**... (M1PE)
- (29) **era** os nossos pais é que **tinham** de responder por nós... (M2PE)
- (30) **é** as senhoras que **estão** aqui (F2PE)

É importante observar que o verbo 'ser' se encontra no singular, mas o verbo que vem depois do complementizador, nas cinco ocorrências, desponta com a marca explícita de plural, mesmo quatro deles (26-29) pertencendo ao grupo de oposição não marcada da escala da saliência fônica, ambiente que desfavorece a concordância verbal.

Além das 22 ocorrências acima mencionadas, aparecem na amostra três construções que se configuram como pseudoclivadas invertidas. São raras no *corpus* e mesmo assim o verbo 'ser' mostra variação⁴⁵.

- (31) os tenderos **é** aqueles que **vendem** a tenda... (M1PE)
 (32) os tenderos **são** aqueles que **vendem** roupas... (M1PE)
 (33) os manajeiros **eram** os que **mandavam** na gente... (FOPE)

4.2.5 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural

Como previsto na hipótese, as chances de ocorrência de marcas de plural são maiores quando ocorre o aumento da saliência do material fônico na oposição singular/plural. Os resultados estão na tabela 12.

Tabela 12: Frequência e peso relativo de concordância verbal conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural no PE.

Fatores	Frequência	PR
Nível 1 – oposição não marcada		
– 1a (<i>vive/vivem</i>)	119/130 = 91,5%	0,299
– 1b (<i>fala/falam</i>)	536/581 = 92,3%	0,380
– 1c (<i>quer/querem</i>)	86/88 = 97,7%	0,427
Nível 2 – oposição marcada		
– 2a (<i>vai/vão</i>)	147/154 = 95,5%	0,477
– 2b (<i>bateu/bateram</i>)	81/86 = 94,2%	0,637
– 2c (<i>teve/tiveram</i>)	371/401 = 92,5%	0,724
Total	1.340/1.440 = 93,1%	Range 425
<i>Input: 0,964</i>		
<i>Significância: 0,015</i>		

⁴⁵ Nesses três casos, consideramos na codificação o SN anteposto ao verbo 'ser'.

Se analisarmos somente as frequências puras, não conseguiremos enxergar com clareza o princípio da saliência fônica. A diferença é notável em termos de peso relativo (*range* de 425). A análise indica que a concordância verbal é desfavorecida com verbos das duas primeiras categorias da oposição não marcada (0,299 e 0,380). Em uma situação próxima ao ponto neutro (0,50), aparecem os verbos da última categoria da oposição não marcada (0,427) e os verbos da primeira categoria da oposição marcada (0,477). E favorecendo mais a concordância, encontramos os verbos das duas últimas categorias da oposição marcada (0,637 e 0,724).

Como não nos deparamos com um salto significativo ao passar do primeiro nível de saliência (0,427), oposição não marcada, para o segundo nível de saliência (0,477), oposição marcada, decidimos verificar a atuação da saliência fônica seguindo a proposta de Lucchesi, Baxter e Silva (2009, p. 350-351). Os resultados estão na tabela 13.

Tabela 13: Frequência e peso relativo de concordância verbal conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural no PE, de acordo com a escala proposta por Lucchesi, Baxter e Silva (2009).

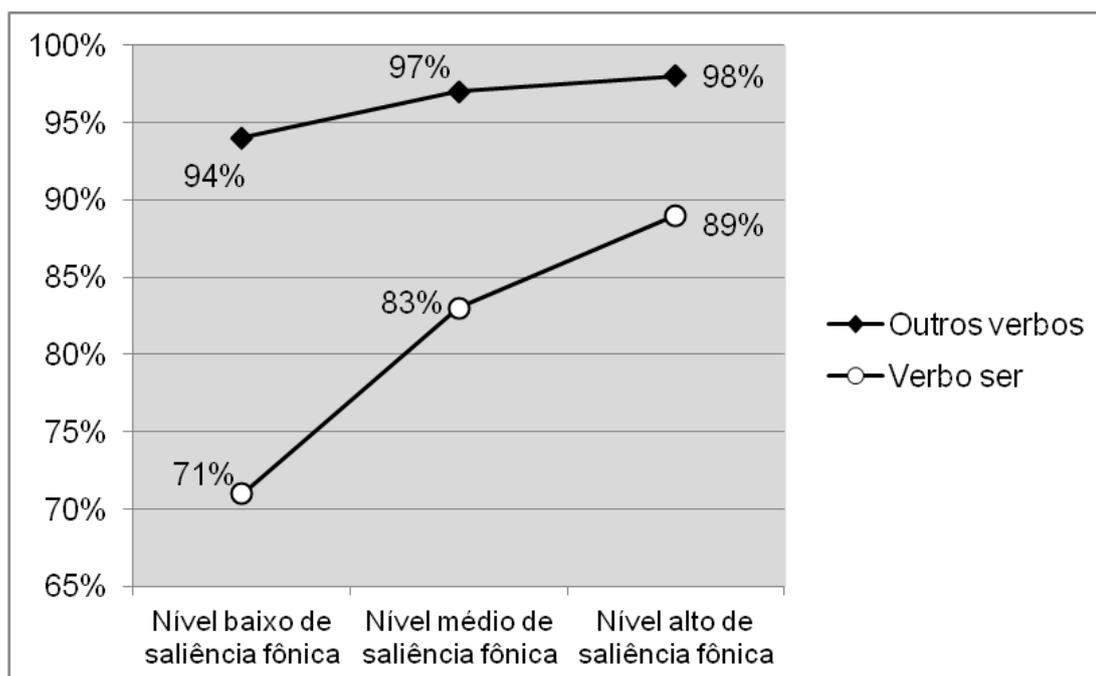
Fatores	Frequência	PR
1. nível baixo de saliência fônica (ex.: <i>sai/saem; bate/batem; fala/falam; era/eram</i>)	655/711 = 92,1%	0,365
2. nível médio de saliência fônica (ex.: <i>faz/fazem; tá/tão; bateu/bateram; quer/querem; vai/vão; foi/foram</i>)	314/328 = 95,7%	0,532
3. nível alto de saliência fônica (ex.: <i>quis/quiseram; fez/fizeram; é/são; veio/vieram</i>)	371/401 = 92,5%	0,706
Total	1.340/1.440 = 93,1%	Range 341
<i>Input: 0,964</i>		
<i>Significância: 0,017</i>		

Na rodada com essa outra configuração da escala, a variável saliência fônica passa a ser selecionada em terceiro lugar pelo programa Goldvarb-X, trocando de posição com a variável tipo estrutural do sujeito/SN, que passa a ser selecionada em quinto lugar. Com as demais variáveis não houve alteração na ordem da seleção estatística.

A tabela 13 aponta, assim como a tabela 12, diferenças estatisticamente significativas em termos de pesos relativos. Aqui o *range* passa a ser de 341.

A fim de observar se há interação da saliência fônica com o tipo de verbo (verbo 'ser' *versus* outros verbos), realizamos o cruzamento dos dois grupos de fatores, cujos resultados vêm expostos no gráfico 5:

Gráfico 5: Frequência de concordância verbal segundo a saliência fônica e o tipo de verbo (verbo 'ser' *versus* outros verbos) no PE.



Com o verbo 'ser' no nível baixo de saliência fônica (*era/eram; fosse/fossem*), a frequência de concordância é de 71% (48/68). Com outros verbos na mesma categoria da saliência fônica (*corre/correm; leva/levam*), a frequência é de 94% (607/643). A diferença de 23 pontos percentuais é bastante significativa.

No nível médio de saliência fônica, ocorrências com o verbo 'ser' (*foi/foram; será/serão*) apresentam 83% (19/23) de frequência de concordância e ocorrências com outros verbos (*está/estão; partiu/partiram*) apresentam 97% (295/305) de frequência. Temos uma diferença de 14 pontos percentuais.

Já no nível alto de saliência fônica, temos o verbo 'ser' (*é/são*) apresentando 89% (224/251) de frequência de concordância e os outros verbos (*começou/começaram*) com um índice de 98% (147/150). Aqui a diferença é de 9 pontos percentuais.

É possível verificar a atuação dos dois grupos de fatores. Há o efeito separado da saliência fônica e o efeito separado do verbo 'ser' e de outros verbos.

Os índices de concordância diminuem conforme a oposição singular/plural dos verbos fica menos saliente. Diminuem ainda mais quando temos o verbo 'ser'. Verificamos que um princípio não anula o outro.

Sobre a atuação da variável saliência fônica no PE, vale informar que não são todas as pesquisas que revelam significância estatística.

O trabalho de Naro e Scherre (2007) e o de Bazenga (2011) confirmam a relevância dessa variável. Os resultados apresentados por Naro e Scherre (2007, p.60) em termos de peso relativo nos oito textos portugueses medievais são: 0,33 com formas verbais menos salientes e 0,75 com formas verbais mais salientes. Varejão (2006) conclui que a saliência fônica da flexão verbal é uma variável relevante, no entanto não apresenta os resultados em termos de pesos relativos.

Na investigação de Monguilhott (2009), não foi selecionada na primeira rodada, mas na segunda rodada, em que exclui os verbos copulativos, a variável passou a ser selecionada como estatisticamente relevante em terceiro lugar. Por fim, na pesquisa de Rubio (2012), não foi selecionada.

4.2.6 Gênero

O último grupo de fatores selecionado pelo programa Goldvarb-X é o gênero. A tabela 14 exhibe os resultados correspondentes para homens e mulheres da nossa amostra do PE.

Tabela 14: Frequência e peso relativo de concordância verbal conforme o gênero no PE.

Fatores	Frequência	PR
– masculino	663/720 = 92,1%	0,430
– feminino	677/720 = 94,0%	0,570
Total	1.340/1.440 = 93,1%	Range 140
<i>Input: 0,964</i>		
<i>Significância: 0,015</i>		

Além de ser o último grupo de fatores selecionado, vemos também que é o grupo que apresenta o menor *range* (140). Isso demonstra que os efeitos linguísticos são mais fortes.

É o primeiro estudo sociolinguístico variacionista com dados do PE que demonstra relevância do gênero. No trabalho de Monguilhott (2009), no de Bazenga (2011) e no de Rubio (2012), a variável gênero não foi selecionada como estatisticamente significativa. Já Varejão (2006) discute somente as variáveis linguísticas.

4.2.7 Distância entre o sujeito/SN e o verbo em número de sílabas⁴⁶

A distância entre o sujeito/SN e o verbo em número de sílabas não obteve significância estatística.

Tabela 15: Frequência de concordância verbal em função do número de sílabas entre o sujeito/SN e o verbo no PE.

Fatores	Frequência
– 0 a 2 sílabas	586/617 = 95,0%
– 3 a 5 sílabas	60/61 = 98,4%
– 6 ou mais sílabas	34/37 = 91,9%
Total	680/715 = 95,1%

Os resultados com as frequências brutas da tabela 15 revelam que as nossas expectativas não foram totalmente confirmadas. Encontramos mais concordância na categoria intermediária com a presença de três a cinco sílabas entre o sujeito/SN e o verbo, tanto que foi encontrado um único dado sem a marca de plural no verbo. Esperávamos uma diminuição progressiva de marca de plural no verbo à medida que o número de sílabas entre o sujeito/SN e o verbo fosse aumentando.

4.2.8 Presença ou ausência do ‘que’ relativo ou complementizador⁴⁷

A presença ou ausência do ‘que’ relativo ou complementizador entre o SN controlador da concordância e o verbo não foi uma variável selecionada como

⁴⁶ No controle desta variável, foram desconsideradas as 526 ocorrências com sujeito nulo e as 199 ocorrências com sujeito/SN posposto ao verbo.

⁴⁷ No controle desta variável, foram desconsideradas as 526 ocorrências com sujeito nulo e as 199 ocorrências com sujeito/SN posposto ao verbo.

estatisticamente relevante pelo programa Goldvarb-X. Os resultados em termos de frequência estão na tabela 16.

Tabela 16: Frequência de concordância verbal em função da presença ou ausência do 'que' relativo ou complementizador no PE.

Fatores	Frequência
– presença do 'que'	171/184 = 92,9%
– ausência do 'que'	509/531 = 95,9%
Total	680/715 = 95,1%

Há uma discreta diferença entre as duas categorias, com a presença do 'que' registrando um pouco menos a marca de plural no verbo.

Naro e Scherre (2007, p. 64) informam que não puderam verificar nos textos antigos portugueses a tendência de o 'que' relativo influenciar a diminuição de marcas quando se coloca entre o SN controlador da concordância e o verbo, e afirmam que “embora esse efeito pareça existir nos dados medievais, tanto em termos de número de dados quanto de pesos relativos, não se obteve significância estatística no nível de 0,05 com os dados disponíveis até o presente momento”.

4.2.9 Paralelismo formal no nível oracional (marcas no sujeito)⁴⁸

O grupo de fatores paralelismo formal no nível oracional também não foi selecionado pelo Goldvarb-X na amostra do PE.

⁴⁸ No controle desta variável, foram desconsideradas as 526 ocorrências com sujeito nulo e as 199 ocorrências com sujeito/SN posposto ao verbo.

Tabela 17: Frequência de concordância verbal segundo o paralelismo formal no nível oracional no PE.

Fatores	Frequência
– presença de plural no último (ou único) elemento	625/656 = 95,3%
– ausência de plural no último elemento	14/16 = 87,5%
– presença de plural no último elemento de um SPrep	9/9 = 100%
– ausência de plural no último elemento de um SPrep	19/19 = 100%
– presença de numeral no último elemento	13/15 = 86,7%
Total	680/715 = 95,1%

Apenas com os índices percentuais expostos na tabela 17, fica difícil argumentar que o paralelismo oracional apresenta as mesmas tendências devidamente comprovadas em amostras do PB.

Essa variável está diretamente ligada à regra de concordância nominal. No PE, o estudo de Brandão e Vieira (2012) evidenciou que a concordância nominal exhibe uma regra quase categórica. As pesquisadoras analisaram 2.449 SNs, presentes na fala de 18 portugueses de Cacém (distribuídos por sexo, três faixas etárias e três níveis de escolaridade), e encontraram 99,96% de aplicação da marca de número. O índice de 0,04% de ausência de marca corresponde a um único SN, constituído por um numeral e um substantivo.

Das 16 ocorrências na nossa amostra sem a marca de plural no último elemento do SN não inserido em um sintagma preposicional, oito são de sujeito composto com o núcleo mais próximo do verbo no singular (34-41). Das oito ocorrências restantes, três aparecem em um contexto de neutralização (42-44) e, de fato, cinco (45-49) sem a marca explícita de plural, sendo que apenas duas (48-49) sem a concordância verbal.

- (34) e agora já tenho uma miúda e:: otro o pai que **vão**... novamente alugá-lo... (F1PE)
- (35) o meu pai os meus tios e o meu avô **vieram** um ano... numa fera (F1PE)
- (36) o Sporting com Benfica e o Porto... **são** rivais entre si... (F1PE)

- (37) olivera... chaparro... ahn::: ... () **são** as árvores principais... do/do Alentejo... (M1PE)
- (38) porque o nome em si e a escritura... **eram** deles não é? (M1PE)
- (39) vamos dizer que o nosso primeiro ministro e o nosso presidente da República... o professor Cavaco Silva... não se **dão** assim lá muito bem... (F2PE)
- (40) o Benfica... o Sporting e o Porto... **são** as equipas assim mais... (F2PE)
- (41) os mais velhos e isso nunca **queriam** aprender... (M2PE)
- (42) e os oto **estudam** (M0PE)
- (43) os oto **estudam** (M0PE)
- (44) as oito hora já **tavam**... marcadas (M0PE)
- (45) os agricultor não **semeavam**... pronto... (M0PE)
- (46) os agricultor já não **davam** trabalho a ninguém... (M0PE)
- (47) os desporto não **gostam** de nós... (F2PE)
- (48) e os lavrador **morava** lá e o meu pai era lá empregado... (F0PE)
- (49) que as pessoa **mata**... é que só no matadoro é que se pode matare... (M0PE)

Destacamos nos nossos dados o uso categórico das marcas de plural nos verbos nos dois contextos com sujeito de estrutura complexa, que apresentam um sintagma preposicional.

Monguilhott (2009) e Rubio (2012) também controlaram este grupo de fatores e, como no nosso estudo, não foi selecionado como estatisticamente relevante.

4.2.10 Escolaridade

Como a variável escolaridade não foi selecionada pelo programa Goldvarb-X, apresentamos na tabela a seguir apenas os resultados percentuais.

Tabela 18: Frequência de concordância verbal de acordo com o grau de escolaridade do informante no PE.

Fatores	Frequência
– não alfabetizados	438/480 = 91,2%
– ensino básico (EJA)	449/480 = 93,5%
– ensino secundário	453/480 = 94,4%
Total	1.340/1.440 = 93,1%

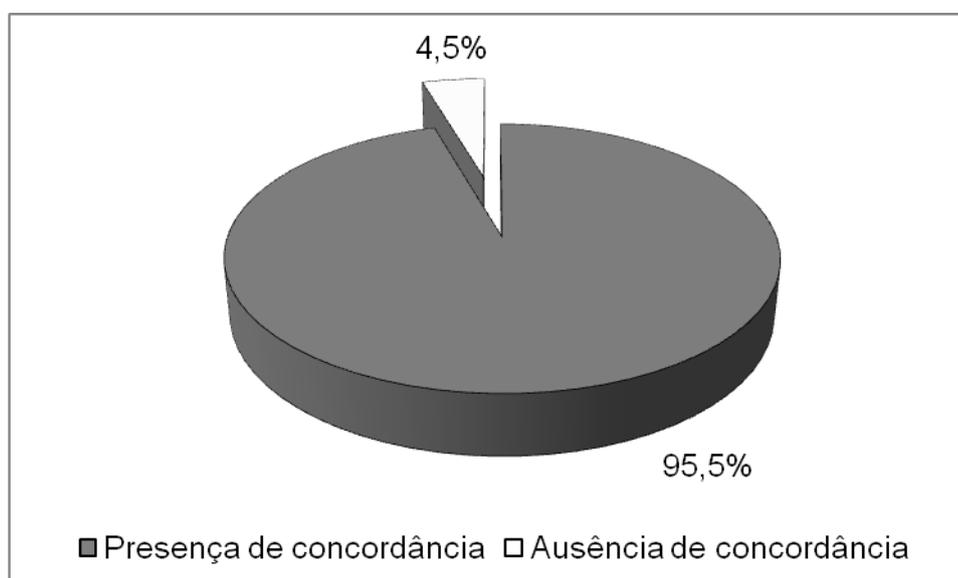
Temos, bem discretamente, um aumento de concordância verbal conforme aumenta o nível de escolaridade. Esses índices poderiam ser interpretados como um indício de que a escolaridade é atuante no PE se levarmos em conta os resultados das pesquisas de Monguilhott (2009) e Bazenga (2011). Na investigação de Monguilhott (2009), foi a única variável social selecionada. Cabe mencionar que a pesquisadora controlou a idade e a escolaridade conjuntamente. Na pesquisa de Bazenga (2011), o nível de escolaridade dos informantes também foi a única variável social significativa.

Não podemos deixar de destacar na nossa amostra o surpreendente índice de frequência de concordância na fala informal de pessoas sem qualquer escolaridade (91,2%). Vemos um alto índice também na fala informal de pessoas com formação supletiva (93,5%). Com 94,4%, é possível afirmar que os estudantes que estavam terminando o ensino secundário apresentam uma regra semicategórica. Retomaremos essa discussão na subseção 4.6 em que comparamos os índices do PB e do PE.

4.3 Análise e discussão dos resultados do PE sem as ocorrências com o verbo 'ser'

Como informamos inicialmente, vamos apresentar os resultados obtidos na amostra do PE sem as ocorrências com o verbo 'ser'. A rodada, composta de 1.098 dados, apresentou apenas 49 ocorrências sem a marca de plural no verbo. Isso significa 4,5% de ausência de concordância, contra 95,5% de presença (1.049/1.098). De acordo com Labov (2003, p. 243), estamos diante de uma regra semicategórica.

Gráfico 6: Distribuição geral dos dados do PE sem as ocorrências com o verbo 'ser'



Com um *input* 0,968 e significância 0,019, os grupos de fatores selecionados na ordem de relevância foram: posição do sujeito/SN em relação ao verbo, traço semântico do sujeito/SN e saliência fônica.

É importante informar que, nessa rodada, os grupos de fatores tipo estrutural do sujeito/SN e presença ou ausência do 'que' relativo ou complementizador não foram selecionados pelo *step-up* e nem excluídos pelo *step-down*. Além disso, a variável posição do sujeito/SN em relação ao verbo foi selecionada pelo *step-up* e, também, excluída pelo *step-down*. Segundo Guy e Zilles (2007, p. 166), "essa situação só ocorre quando se trata de uma análise complexa (com muitos grupos de

fatores), e quando os grupos não são completamente ortogonais, em termos da distribuição de dados”.

4.3.1 Posição do sujeito/SN em relação ao verbo⁴⁹

Os resultados da posição do sujeito/SN em relação ao verbo, sem os dados com o verbo ‘ser’, podem ser vistos na tabela 19:

Tabela 19: Frequência e peso relativo de concordância verbal segundo a posição do sujeito/SN em relação ao verbo no PE, sem as ocorrências com o verbo ‘ser’.

Fatores	Frequência	PR
– sujeito/SN anteposto	538/555 = 96,9%	0,560
– sujeito/SN posposto	67/84 = 79,8%	0,169
Total	605/639 = 94,7%	<i>Range</i> 391
<i>Input</i> : 0,968		
Significância: 0,019		

As chances de marca explícita de plural no verbo diminuem com o uso do sujeito/SN posposto (79,8%). Como esperado, as chances aumentam com o sujeito/SN anteposto (96,9%). A diferença entre os dois fatores é de 17,1 pontos percentuais e um *range* de 391.

4.3.2 Traço semântico do sujeito/SN

Nossa análise evidencia que o traço [+ humano] desempenha um papel importante na concordância verbal no PE.

⁴⁹ No controle desta variável, foram desconsideradas as 459 ocorrências com sujeito nulo.

Tabela 20: Frequência e peso relativo de concordância verbal em função do traço semântico do sujeito/SN no PE, sem as ocorrências com o verbo 'ser'.

Fatores	Frequência	PR
– SN [+ humano / + animado]	914/944 = 96,8%	0,536
– SN [– humano / + animado]	19/20 = 95,0%	0,383
– SN [– humano / – animado]	116/134 = 86,6%	0,280
Total	1.049/1.098 = 95,5%	Range 256
<i>Input: 0,968</i>		
<i>Significância: 0,019</i>		

Nos contextos de sujeito/SN [+ humano], o verbo apresenta maior possibilidade de vir no plural (0,536). Se o sujeito/SN for [– humano], a possibilidade de vir no plural é menor (0,383 / 0,280).

4.3.3 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural

A saliência fônica, mais uma vez, se mostra atuante. As formas verbais com baixo nível de saliência são mais propensas ao cancelamento da marca de número (peso relativo 0,405). Por outro lado, as formas verbais mais salientes favorecem a concordância. No nível médio de saliência, o peso relativo é de 0,597 e no nível alto de saliência é de 0,698.

Tabela 21: Frequência e peso relativo de concordância verbal conforme o grau de saliência fônica no PE, sem as ocorrências com o verbo 'ser'.

Fatores	Frequência	PR
1. nível baixo de saliência fônica (ex.: <i>sai/saem; bate/batem; fala/falam; ganha/ganham</i>)	607/643 = 94,4%	0,405
2. nível médio de saliência fônica (ex.: <i>faz/fazem; tá/tão; bateu/bateram; quer/querem; vai/vão</i>)	295/305 = 96,7%	0,597
3. nível alto de saliência fônica (ex.: <i>quis/quiseram; fez/fizeram; veio/vieram</i>)	147/150 = 98,0%	0,698
Total	1.049/1.098 = 95,5%	Range 293
<i>Input: 0,968</i>		
<i>Significância: 0,019</i>		

Como podemos ver na tabela 21, o maior número de ocorrências sem a marca de plural no verbo se concentra no nível mais baixo de saliência fônica (36 dados). O fator nível médio traz 10 dados com o cancelamento da marca e, por sua vez, o fator nível alto de saliência fônica com apenas três dados (50-52) de ausência.

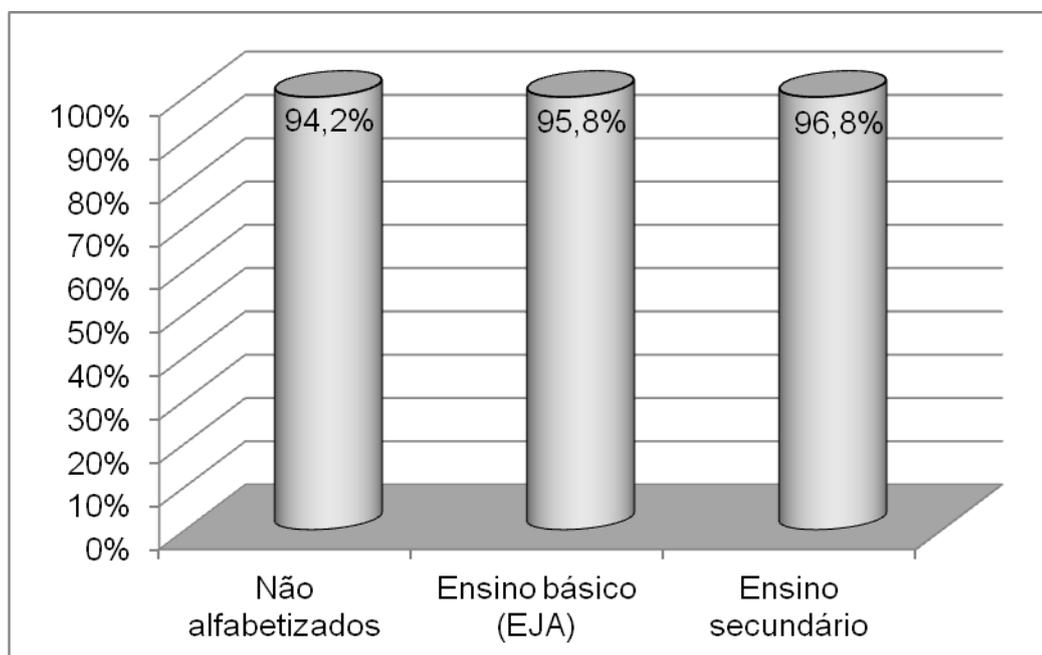
- (50) **começô** os anos a passare... (F0PE)
- (51) quando isso ahn/ **começô** aparecê mais (essas) ceferas... (M0PE)
- (52) já tem se feito (algumas) greves mas nunca **deu**... nunca **deu** resultado... (M2PE)

As três ocorrências possuem o sujeito inanimado, e em duas (50-51) o sujeito está posposto ao verbo. Como já vimos, sujeito/SN posposto e com traços [- humano] e [- animado] favorece o cancelamento da marca de plural no verbo.

4.3.4 Escolaridade

A variável escolaridade continuou não sendo selecionada pelo programa Goldvarb-X. Apresentamos os resultados para mostrar que, sem as ocorrências com o verbo 'ser', o *status* semicategórico da regra na amostra do PE fica mais evidente.

Gráfico 7: Frequência de concordância verbal de acordo com o grau de escolaridade do informante no PE, sem as ocorrências com o verbo 'ser'.

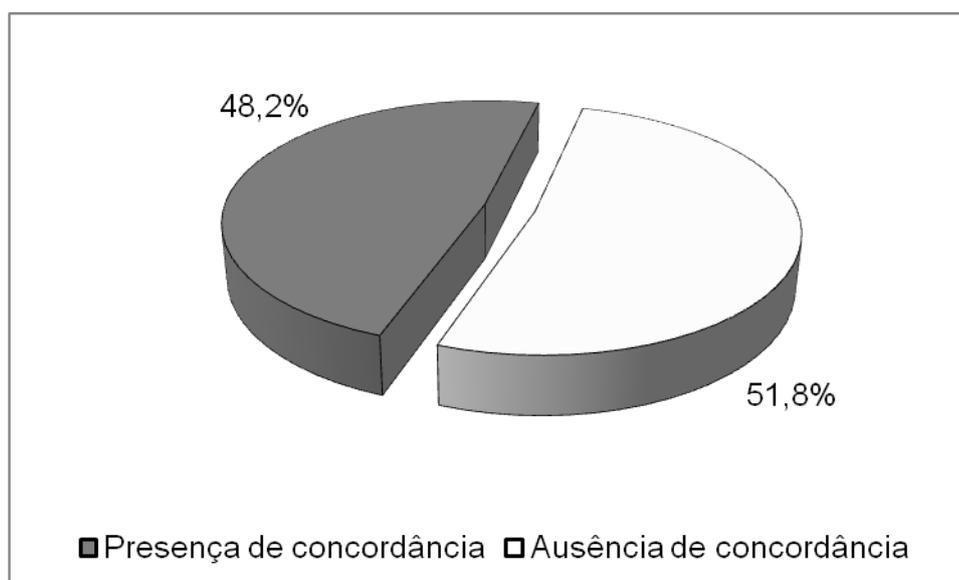


As diferenças entre os níveis de escolaridade não são estatisticamente relevantes. Os informantes do ensino secundário apresentam um índice um pouco maior de concordância verbal ($328/339 = 96,8\%$), com apenas 11 dados sem a marca formal de plural nos verbos. Os informantes do ensino básico com formação supletiva apresentam 95,8% ($346/361$) de concordância verbal, com 15 dados de cancelamento da marca de plural nos verbos. Depois vem o grupo de pessoas não alfabetizadas com uma frequência de 94,2% ($375/398$) de aplicação da regra de concordância, com 23 dados sem a marca explícita de plural nos verbos.

4.4 Análise e discussão dos resultados gerais do PB

A amostra de São Carlos (PB) é composta de 1.422 ocorrências de terceira pessoa do plural, com 686 ocorrências (48,2%) apresentando a marca formal de plural nos verbos e 736 (51,8%) ocorrências sem a marca formal de plural nos verbos.

Gráfico 8: Distribuição geral dos dados do PB.



A presença da concordância verbal apresentou *input* 0,466 e a rodada significância 0,042. O programa Goldvarb-X selecionou, na ordem de relevância, os seguintes grupos de fatores: escolaridade, saliência fônica, tipo estrutural do sujeito/SN, paralelismo formal no nível oracional, gênero, posição do sujeito/SN em relação ao verbo, traço semântico do sujeito/SN, tipo de verbo (verbo 'ser' *versus* outros verbos) e distância entre o sujeito/SN e o verbo em número de sílabas.

A variável presença ou ausência do 'que' relativo ou complementizador não foi selecionada pelo *step-up* e nem excluída pelo *step-down*. Já os grupos de fatores tipo estrutural do sujeito/SN e tipo de verbo (verbo 'ser' *versus* outros verbos) foram selecionados pelo *step-up* e também excluídos pelo *step-down*. No entanto, na rodada com a configuração da escala da saliência fônica em três níveis (nível baixo, médio e alto de saliência), as variáveis presença ou ausência do 'que' relativo ou complementizador e tipo de verbo (verbo 'ser' *versus* outros verbos) foram

selecionadas normalmente. Apenas o grupo de fatores tipo estrutural do sujeito/SN continuou sendo selecionado pelo *step-up* e excluído pelo *step-down*.

Sendo assim, é possível confirmar que todos os grupos de fatores analisados apresentam significância estatística na nossa amostra com dados do PB.

Apresentaremos, em primeiro lugar, os resultados e as discussões das variáveis linguísticas e, por último, das variáveis sociais.

4.4.1 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural

O segundo grupo de fatores selecionado pelo programa Goldvarb-X foi a saliência fônica. Os resultados confirmam que o crescente índice de concordância está correlacionado com a crescente saliência fônica da oposição singular/plural dos verbos.

Tabela 22: Frequência e peso relativo de concordância verbal conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural no PB.

Fatores	Frequência	PR
Nível 1 – oposição não marcada		
– 1a (<i>vive/vivem</i>)	30/89 = 33,7%	0,142
– 1b (<i>fala/falam</i>)	186/613 = 30,3%	0,224
– 1c (<i>quer/querem</i>)	21/55 = 38,2%	0,213
Nível 2 – oposição marcada		
– 2a (<i>vai/vão</i>)	117/187 = 62,6%	0,723
– 2b (<i>bateu/bateram</i>)	73/125 = 58,4%	0,805
– 2c (<i>teve/tiveram</i>)	259/353 = 73,4%	0,858
Total	686/1.422 = 48,2%	Range 716
<i>Input</i> : 0,466		
Significância: 0,042		

Como verificamos na tabela acima, os verbos das categorias menos salientes (nível 1) inibem a concordância verbal. Já os verbos das categorias mais salientes (nível 2) favorecem o uso da concordância.

O maior salto no índice de concordância ocorre entre os níveis 1 e 2 da hierarquia oposicional. É bastante evidente essa divisão quando observamos os pesos relativos associados aos fatores. Esse salto é de 20,2 pontos percentuais em

termos de frequência e 510 em termos de peso relativo. Não podemos deixar de destacar o alto *range* dessa variável (716).

Na tabela 23, a seguir, mostramos os resultados obtidos para a variável saliência fônica com a escala proposta por Lucchesi, Baxter e Silva (2009, p. 350-351), assim como fizemos com os dados do PE.

Tabela 23: Frequência e peso relativo de concordância verbal conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural no PB, de acordo com a escala proposta por Lucchesi, Baxter e Silva (2009).

Fatores	Frequência	PR
1. nível baixo de saliência fônica (ex.: <i>sai/saem; bate/batem; fala/falam; era/eram</i>)	216/702 = 30,8%	0,223
2. nível médio de saliência fônica (ex.: <i>faz/fazem; tá/tão; bateu/bateram; quer/querem; vai/vão; foi/foram</i>)	211/367 = 57,5%	0,677
3. nível alto de saliência fônica (ex.: <i>quis/quiseram; fez/fizeram; é/são; veio/vieram</i>)	259/353 = 73,4%	0,848
Total	686/1.422 = 48,2%	Range 625
<i>Input: 0,464</i>		
<i>Significância: 0,046</i>		

Verificamos o aumento na frequência de concordância obedecendo a hierarquia da escala da saliência fônica. Em termos de frequência, a diferença entre o nível mais baixo e o nível mais alto de saliência fônica é de 42,6 pontos percentuais. O *range* de 625 continua sendo forte.

A título de comparação, trazemos os resultados obtidos por Lucchesi, Baxter e Silva (2009, p. 350-351), confrontados aos nossos.

Tabela 24: frequência e peso relativo de concordância verbal conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural, no *corpus* sob análise e no *corpus* de Lucchesi, Baxter e Silva (2009).

Fatores	Frequência de concordância verbal			
	Nossos resultados		Lucchesi, Baxter e Silva (2009)	
	Frequência	PR	Frequência	PR
1. nível baixo de saliência fônica (ex.: <i>sai/saem; bate/batem; fala/falam; era/eram</i>)	216/702 = 30,8%	0,223	48/818 = 6%	0,27
2. nível médio de saliência fônica (ex.: <i>faz/fazem; tá/tão; bateu/bateram; quer/querem; vai/vão; foi/foram</i>)	211/367 = 57,5%	0,677	158/675 = 23%	0,69
3. nível alto de saliência fônica (ex.: <i>quis/quiseram; fez/fizeram; é/são; veio/vieram</i>)	259/353 = 73,4%	0,848	67/213 = 31%	0,78

Como podemos observar na tabela 24, os resultados em termos de pesos relativos nas duas pesquisas são muito semelhantes: o nível baixo da hierarquia da saliência favorece mais a ausência de concordância do que os níveis médio e alto.

Vale lembrar que Lucchesi, Baxter e Silva (2009) trabalharam com falantes com pouca ou nenhuma escolarização das comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia. Na nossa amostra, há jovens com o ensino médio, o que faz com que esses resultados sejam notáveis, pois os percentuais altos de concordância se devem à escolaridade mais alta desse grupo do nosso *corpus*.

O princípio da saliência fônica tem sido aplicado em todos os trabalhos sociolinguísticos sobre a variação da concordância verbal de 3ª pessoa do plural no PB. Os resultados aqui expostos confirmam a hipótese formulada por Lemle e Naro (1977) e vêm corroborar os de outros estudos (cf. NARO, 1981; RODRIGUES, 1987; VIEIRA, 1995; NARO e SCHERRE, 1999a; MONGUILHOTT, 2001; PEREIRA, 2004; GAMEIRO, 2005; MONTE, 2007; RUBIO, 2008; entre outros) que também controlaram este grupo de fatores.

4.4.2 Tipo estrutural do sujeito/SN

Na tabela 25 estão expostos os resultados obtidos para o grupo de fatores tipo estrutural do sujeito/SN, que foi selecionado em terceiro lugar pelo programa Goldvarb-X⁵⁰.

Tabela 25: Frequência e peso relativo de concordância verbal de acordo com o tipo estrutural do sujeito/SN no PB.

Fatores	Frequência	PR
– Numeral (núcleo ou adjunto)	37/125 = 29,6%	0,660
– Pronome <i>vocês</i>	14/19 = 73,7%	0,583
– Pronome pessoal <i>eles/elas</i>	311/503 = 61,8%	0,565
– SN pleno nu	14/22 = 63,6%	0,540
– Sujeito nulo	162/350 = 46,3%	0,500
– SN pleno simples	95/284 = 33,5%	0,400
– SN pleno composto	11/26 = 42,3%	0,309
– Quantificador indefinido	39/83 = 47,0%	0,306
– Pronome demonstrativo	3/10 = 30,0%	0,123
Total	686/1.422 = 48,2%	Range 537
<i>Input: 0,466</i>		
<i>Significância: 0,042</i>		

Desperta a atenção o fator com numeral, já que apresenta o menor índice de frequência de concordância (29,6%) e o peso relativo mais forte (0,660). Vale lembrar que não consideramos nesse fator apenas o numeral como núcleo. Em muitos casos, o numeral é o adjunto e, dessa forma, reforça mais a pluralidade no sintagma em que está incluído.

Confirmando a hipótese, encontramos o pronome ‘vocês’ e o pronome pessoal ‘eles/elas’ favorecendo a marcação da concordância verbal. Com esses pronomes, a posposição, fator que inibe a concordância, quase não ocorre. O pronome ‘vocês’ aparece em 19 ocorrências e em todas está anteposto ao verbo. O maior número de dados da nossa amostra está no fator com o pronome pessoal ‘eles/elas’. São 503 dados, sendo que apenas três trazem o pronome ‘eles/elas’ posposto ao verbo (53-55), que não apresentam, então, marcas de concordância:

⁵⁰ Lembramos que esse grupo de fatores foi selecionado pelo *step-up* (3º lugar) e, também, excluído pelo *step-down*.

- (53) **era** novo eles... (M0PB)
 (54) não **era** elas ainda... eu nem sei quem que era a diretora (M1PB)
 (55) **é** sempre eles que... que fazem a maioria dos shows... (F2PB)

Ainda favorecendo a concordância, encontramos o SN pleno nu (63,6% de frequência e 0,540 de peso relativo). Um dado importante é que não houve ocorrências de SN pleno nu na fala das pessoas não alfabetizadas.

O sujeito nulo se encontra em uma faixa intermediária. Nossa expectativa era a de que favoreceria mais a presença de concordância, pois, de acordo com Rodrigues (1989, p. 550-551), “é lícito pensarmos num hipotético comprometimento do conteúdo informacional da frase já que, isoladamente, a forma verbal não-marcada não indica formal e semanticamente o sujeito da oração”. Entretanto, analisando os nossos dados, chegamos à mesma conclusão da pesquisadora:

[...] a ausência simultânea de sujeito pronominal e marcas de pluralidade no verbo não constitui obstáculo para a apreensão do seu sujeito, já que ele está fisicamente presente em orações que antecedem imediatamente as formas verbais em questão. (RODRIGUES, 1989, p. 552).

Os sujeitos/SNs pleno simples (0,400 de peso relativo), pleno composto (0,309 de peso relativo), quantificador indefinido (0,306 de peso relativo) e pronome demonstrativo (0,123 de peso relativo) foram os que menos preservaram as marcas de concordância verbal, respectivamente.

Apresentando o menor peso relativo, a categoria com pronome demonstrativo exibe poucos dados. Vejamos:

- (56) aqueles dali **acompanha**... (F0PB)
 (57) pos **que tava** já (tá) trabalhano... (F0PB)
 (58) os **que aprendeu** mais os mai novo né?... (M0PB)
 (59) os **que fica**... vi/ vigia lá na/ nas porta... (M0PB)
 (60) então dá pa conta os **que tá** lá faz tempo (F1PB)
 (61) mas aqueles que a/ que a mãe... falô essas coisas de Deus... **pende** pa esse lado (F1PB)
 (62) os **que tão** lá... (F1PB)
 (63) e os **que ficaram** também... (M2PB)
 (64) **é** aqueles também **que querem** né? ajuda (F2PB)

Dos 10 dados, sete são de pessoas com nula ou pouca escolaridade (56-62), e na maioria dos casos encontramos o ‘que’ entre o pronome demonstrativo controlador da concordância e o verbo. Esses fatores influenciam o cancelamento da marca de plural nos verbos.

Devemos registrar que houve um número pequeno de ocorrências em algumas categorias e que se trata de um grupo de fatores complexo, com interferência de outras variáveis.

4.4.3 Paralelismo formal no nível oracional (marcas no sujeito)⁵¹

O quarto grupo de fatores selecionado estatisticamente foi o paralelismo formal no nível oracional, cujos resultados são apresentados na tabela 26.

Tabela 26: Frequência e peso relativo de concordância verbal segundo o paralelismo formal no nível oracional no PB.

Fatores	Frequência	PR
– presença de plural no último (ou único) elemento	434/688 = 63,1%	0,625
– ausência de plural no último elemento	28/178 = 15,7%	0,175
– presença de plural no último elemento de um SPrep	1/1 = 100%	–
– ausência de plural no último elemento de um SPrep	7/17 = 41,2%	0,395
– presença de numeral no último elemento	11/28 = 39,3%	0,082
Total	481/912 = 52,7%	Range 543
<i>Input: 0,466</i>		
<i>Significância: 0,042</i>		

Como era esperado, há realmente maior frequência de realização de concordância verbal quando a marca formal de plural está presente no último ou único elemento do sujeito. Essa tendência fica mais evidente ao analisarmos os pesos relativos. O *range* de 543 mostra a importância dessa variável na análise estatística realizada.

⁵¹ No controle desta variável, foram desconsideradas as 350 ocorrências com sujeito nulo e as 160 ocorrências com sujeito/SN posposto ao verbo.

Lembramos que nossa hipótese era justamente essa: “marcas conduzem a marcas e zeros conduzem a zeros” (SCHERRE e NARO, 1993, p. 11).

Encontramos um número maior de ocorrências nos dois fatores em que o sujeito não inclui em sua estrutura um complemento na forma de sintagma preposicional (SPrep). Dos 688 dados com a presença de plural no último (ou único) elemento não inserido em um SPrep, 434 apresentam a marca de plural nos verbos (63,1% de frequência e 0,625 de peso relativo). Dos 178 dados com a presença da forma de plural zero no último elemento não inserido em um SPrep (conforme exemplos (65-68)), apenas 28 trazem a marca de plural nos verbos (15,7% de frequência e 0,175 de peso relativo). Entre essas duas categorias, temos uma diferença de 47,4% pontos percentuais em termos de frequência e 450 em termos de peso relativo.

(65) os menino_ **fica** aqui comigo né? (F0PB)

(66) minhas filha_ **nasceu** aqui memo (M0PB)

(67) os filho_ **tá** na escola (F1PB)

(68) os evangélico_ **é** assim (M1PB)

Em uma circunstância intermediária, mas ainda desfavorecendo a marca de plural nos verbos, encontramos a categoria com a presença da forma zero no último elemento do SPrep. São poucos dados nessa categoria, mas a tendência se confirma, apresentando 41,2% (7/17) de frequência e 0,395 de peso relativo.

Com a presença da forma de plural explícita no último elemento inserido em um SPrep, obtivemos uma única ocorrência no nosso *corpus* do PB, que apresenta, então, a marca explícita de plural no verbo.

Em relação à presença de numeral no último elemento, podemos verificar que esse fator favorece o cancelamento da marca de plural nos verbos (com peso relativo 0,082). Analisando somente os índices de frequência, não vemos uma diferença muito significativa entre a categoria com a presença da forma zero no último elemento do SPrep (41,2%) e a categoria com a presença de numeral no último elemento (39,3%). Entretanto, os pesos relativos revelam um comportamento diferente. Por esse motivo, resolvemos analisar as 17 ocorrências (69-85) em que não há marca explícita de plural nos verbos em que o último elemento do sujeito é um numeral.

- (69) às veze tinha duas que **tava** dano leite (F0PB)
- (70) tem só dois que **faleceu**... (F0PB)
- (71) dois que **morreu** (M0PB)
- (72) dois já **tá** terminando... (F1PB)
- (73) dois **trabalha** na secretaria (F1PB)
- (74) dois **trabalha** de pedrero (M1PB)
- (75) as duas **funciona** normal... (M1PB)
- (76) tem dois que **toca** na igreja né? (M1PB)
- (77) tem dois que **toca** violão... (M1PB)
- (78) por parte de pai os dois já **faleceu**... (F2PB)
- (79) os doi já se **foi**... (M0PB)
- (80) os doi **tinha** poco escola (M0PB)
- (81) no Cidade Aracy... três quatro por noite... **morria**... (M1PB)
- (82) quatro ou cinco que **vai** lá fazê (M1PB)
- (83) teve uns trei ou quatro lá que **bombô**... tanta falta... (M1PB)
- (84) os cinco **mora** comigo... (M1PB)
- (85) oito **levantô** a mão... (F2PB)

Das 17 ocorrências, 15 são de falantes de nula ou pouca escolaridade, sendo que 11 dados são de pessoas do sexo masculino. Como veremos adiante, a escolaridade e o gênero são determinantes na variação da concordância verbal no PB. Cinco ocorrências (81-85) são de numerais não terminados em 's' e em duas (79-80) o falante não realiza o 's' final do numeral 'dois'. Além disso, nove dados apresentam formas verbais com baixa saliência fônica.

4.4.4 Posição do sujeito/SN em relação ao verbo⁵²

Selecionada em sexto lugar, a análise confirmou a relevância estatística da variável posição do sujeito/SN em relação ao verbo. Os resultados dessa variável podem ser encontrados na tabela 27, a seguir.

Tabela 27: Frequência e peso relativo de concordância verbal segundo a posição do sujeito/SN em relação ao verbo no PB.

Fatores	Frequência	PR
– sujeito/SN anteposto	481/912 = 52,7%	0,613
– sujeito/SN posposto	43/160 = 26,9%	0,068
Total	524/1.072 = 48,9%	Range 545
<i>Input: 0,466</i>		
<i>Significância: 0,042</i>		

Tal como era esperado, os sujeitos/SNs pospostos ao verbo tendem a enfraquecer consideravelmente a concordância verbal, com 26,9% de frequência e 0,068 de peso relativo. A categoria mais saliente, a do sujeito/SN anteposto, favorece a aplicação da regra de concordância, com 52,7% de frequência e 0,613 de peso relativo. O *range* de 545 revela o efeito significativo da variável posicional.

Uma das propriedades sintáticas consideradas características dos verbos inacusativos é a posposição do sujeito/SN. Na nossa amostra, muitos dados de SN posposto parecem se caracterizar por verbos que podem ser analisados como inacusativos ou por construções apresentativas. Vejamos alguns exemplos:

- (86) aí quando **chegô** as máquina miorô um poco né?... (M0PB)
- (87) **morreu** treis irmão e/ e uma irmã minha... (M0PB)
- (88) **chegava** aqueles caminhão carregado de adubo... (M1PB)
- (89) **saiu** umas coisa no jornal sabe? (F1PB)
- (90) **existem** dois tipos no mundo da moda (M2PB)
- (91) **era::: era** aquelas injeção (F0PB)
- (92) **era** pocas casa que tinha aqui (M1PB)
- (93) **são...** quatro acho... quatro ou cinco... (F2PB)

⁵² No controle desta variável, foram desconsideradas as 350 ocorrências com sujeito nulo.

Sobre a posposição do SN, Bagno (2011), em sua gramática pedagógica do português brasileiro, propõe a seguinte regra:

Assim, como já afirmamos também sobre os inacusativos, as duas análises são igualmente válidas: tanto a que considera uma estrutura Verbo–Sujeito e, portanto, prevê a concordância, quanto a que considera uma estrutura Verbo impessoal–Sintagma nominal no caso absoluto e, portanto não prevê a concordância. (BAGNO, 2011, p.655).

4.4.5 Traço semântico do sujeito/SN

Pelos resultados da tabela 28, constatamos que o traço semântico do sujeito/SN atua no condicionamento da concordância verbal. A variável foi selecionada em sétimo lugar pelo programa Goldvarb-X.

Tabela 28: Frequência e peso relativo de concordância verbal em função do traço semântico do sujeito/SN no PB.

Fatores	Frequência	PR
– SN [+ humano / + animado]	656/1.284 = 51,1%	0,522
– SN [– humano / + animado]	0/24 = 0,0%	–
– SN [– humano / – animado]	30/114 = 26,3%	0,274
Total	686/1.422 = 48,2%	Range 248
<i>Input: 0,466</i>		
<i>Significância: 0,042</i>		

A hipótese foi confirmada parcialmente, pois esperávamos um valor intermediário da aplicação da regra nos contextos com sujeito/SN [– humano / + animado]. Nos 24 casos dessa categoria, o cancelamento da marca de plural foi categórico, conforme os exemplos abaixo:

- (94) as abeia que **mordia** nós... (F0PB)
- (95) porque senão elas **bate** o pé né? (elas = vacas) (F0PB)
- (96) até os bezerro grande **comia**... (M0PB)
- (97) os bezerro **tava** grande... (M0PB)

Voltando aos dados, pudemos verificar que das 24 ocorrências, 22 foram pronunciadas por informantes não alfabetizados e duas por um homem de escolaridade supletiva (EJA). Não houve qualquer dado com sujeito/SN [– humano / + animado] no grupo dos jovens com mais escolaridade (ensino médio). Um tema recorrente nas entrevistas sociolinguísticas com os informantes mais velhos foi a vida no campo, porque muitos tiveram uma relação forte com a zona rural. É aí que aparecem os animais, que não são humanos, mas são animados. Portanto, uma explicação plausível é a forte atuação da escolaridade nessa categoria.

O que de fato se mostra relevante é a influência do traço semântico [+humano], como em (98-99), favorecendo a aplicação da regra de concordância, e o traço semântico [– humano], como em (100-101), inibindo a concordância verbal.

- (98) e muitos traficante **tão** até hoje aí né?... (F1PB)
 (99) oito amigos... quase todos **eram** assim... (M2PB)
 (100) algumas coisa que **tá**... faltando né? (M1PB)
 (101) aquelas tampinha terminal... que **vai** no compresor... (M1PB)

4.4.6 Tipo de verbo: verbo ‘ser’ versus outros verbos

Na ordem de relevância, o tipo de verbo (verbo ‘ser’ versus outros verbos) foi o oitavo selecionado e apresenta o menor *range* (157) comparado ao dos outros grupos de fatores da nossa análise.

Tabela 29: Frequência e peso relativo de concordância verbal segundo o tipo de verbo (verbo ‘ser’ versus outros verbos) no PB.

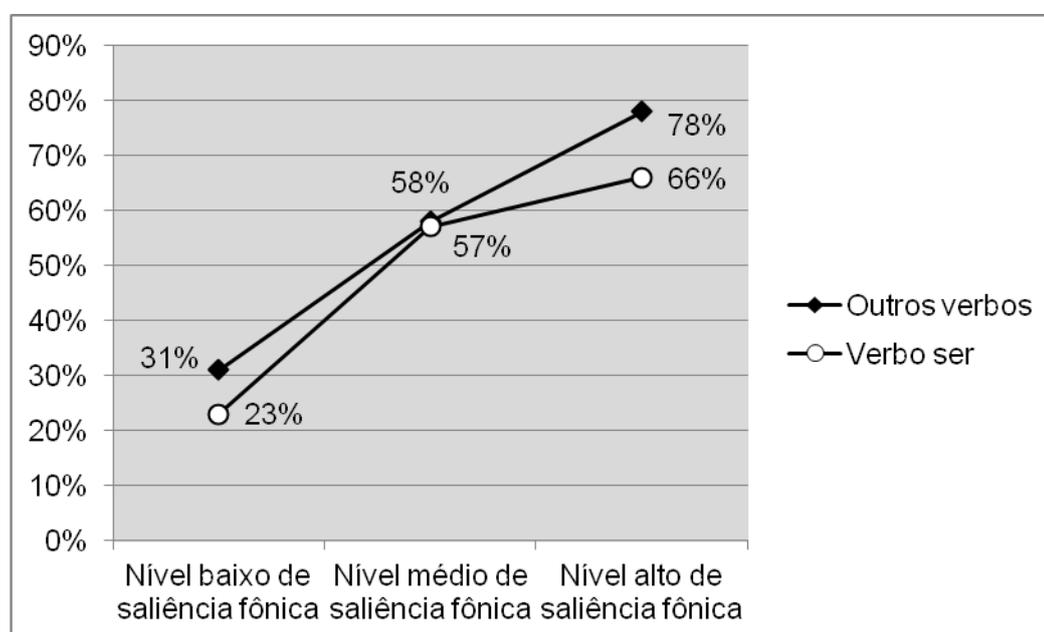
Fatores	Frequência	PR
– verbo ‘ser’	118/218 = 54,1%	0,632
– outros verbos	568/1.204 = 47,2%	0,475
Total	686/1.422 = 48,2%	Range 157
<i>Input</i> : 0,466		
Significância: 0,042		

Com os dados da nossa amostra do PB, a hipótese não foi confirmada. Esperávamos um maior índice de cancelamento da marca de plural em construções

com o verbo 'ser', mas os valores de frequência e pesos relativos da tabela 29 mostram o contrário. Com uma diferença não muito grande, o verbo 'ser' aparece favorecendo a concordância (54,1% de frequência e 0,632 de peso relativo) em relação aos outros verbos (47,2% de frequência e 0,475 de peso relativo).

A fim de verificar melhor o comportamento do verbo 'ser' na amostra do PB, realizamos o cruzamento com a variável saliência fônica verbal. O gráfico 9 mostra os resultados desse cruzamento.

Gráfico 9: Frequência de concordância verbal segundo a saliência fônica e o tipo de verbo (verbo 'ser' versus outros verbos) no PB.



Como podemos observar, a saliência fônica é mais atuante no condicionamento da regra de concordância verbal. Os índices de concordância aumentam conforme a oposição singular/plural dos verbos fica mais saliente.

No nível baixo de saliência fônica, ocorrências com o verbo 'ser' (*era/eram; fosse/fossem*) apresentam 23% (13/56) de frequência de concordância e ocorrências com outros verbos (*corre/correm; leva/levam*) apresentam 31% (203/646) de frequência. É uma diferença de apenas 8 pontos percentuais.

Com o verbo 'ser' no nível médio de saliência fônica (*foi/foram; será/serão*), a frequência de concordância é de 57% (12/21). Com outros verbos na mesma categoria da saliência fônica (*está/estão; partiu/partiram*), a frequência é de 58% (199/346). Houve um empate em termos percentuais.

Já no nível alto de saliência fônica, temos o verbo 'ser' (é/são) apresentando 66% (93/141) de frequência de concordância e os outros verbos (começou/começaram) com um índice de 78% (166/212). Temos uma diferença de 12 pontos percentuais.

Sabemos que o verbo 'ser' ocorre com grande frequência nas sentenças apresentacionais e, nessas sentenças, aparece antes do SN que ele introduz (cf. exemplos 102-105).

(102) Doc.: e quem trabalhava?... Inf.: **eram** pessoas voluntárias (F1PB)

(103) **era** aqueles banco de madeira... (M1PB)

(104) **são** três irmãs e ele tem dó... (F1PB)

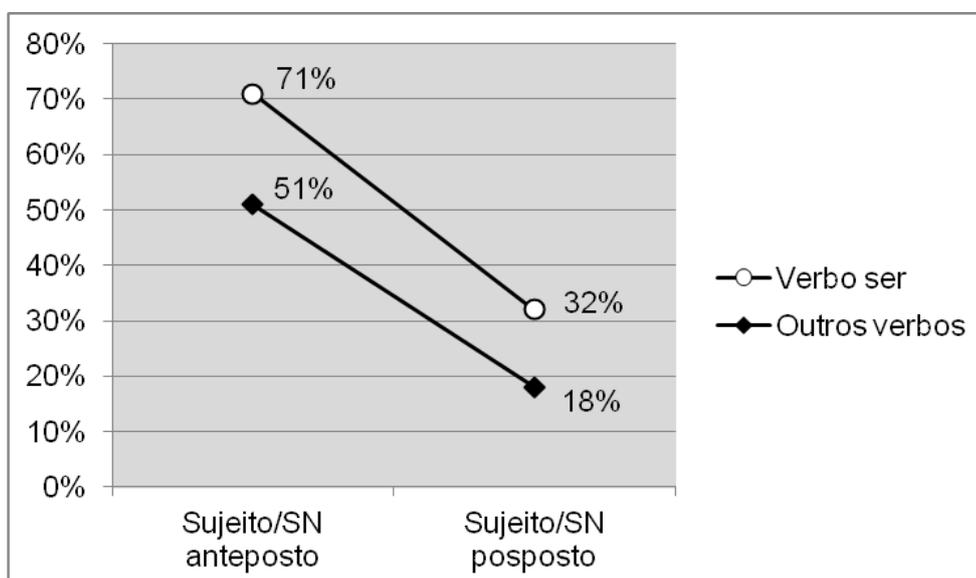
(105) acho que **é** mais os amigos... com certeza... (M2PB)

Sendo assim, realizamos o cruzamento do tipo de verbo (verbo 'ser' *versus* outros verbos) com a posição do sujeito/SN. Vejamos os resultados encontrados:

Tabela 30: Frequência de concordância verbal segundo o tipo de verbo (verbo 'ser' *versus* outros verbos) e a posição do sujeito/SN em relação ao verbo no PB.

Fatores	Frequência de concordância	
	– sujeito/SN anteposto	– sujeito/SN posposto
– verbo 'ser'	55/78 = 71%	33/104 = 32%
– outros verbos	426/834 = 51%	10/56 = 18%

Gráfico 10: Frequência de concordância verbal segundo o tipo de verbo (verbo 'ser' *versus* outros verbos) e a posição do sujeito/SN em relação ao verbo no PB.



Observando os resultados acima, vemos que a frequência de concordância decresce, tanto com o verbo 'ser' quanto com outros verbos, quando o sujeito/SN está posposto. Mas, diferentemente dos resultados encontrados na amostra do PE (cf. tabela 11 e gráfico 4), os índices de frequência de concordância são maiores com o verbo 'ser'.

4.4.7 Distância entre o sujeito/SN e o verbo em número de sílabas⁵³

A nona variável selecionada foi a distância entre o sujeito/SN e o verbo em número de sílabas. Os valores são apresentados na tabela 31:

Tabela 31: Frequência e peso relativo de concordância verbal em função do número de sílabas entre o sujeito/SN e o verbo no PB.

Fatores	Frequência	PR
– 0 a 2 sílabas	459/854 = 53,7%	0,515
– 3 a 5 sílabas	19/44 = 43,2%	0,398
– 6 ou mais sílabas	3/14 = 21,4%	0,094
Total	481/912 = 52,7%	<i>Range 421</i>
<i>Input: 0,466</i>		
<i>Significância: 0,042</i>		

Há uma diminuição na frequência de concordância à medida que o sujeito/SN vai se distanciando do verbo, confirmando a hipótese estabelecida. O princípio da saliência posicional também está presente nesse grupo de fatores (cf. LEMLE e NARO, 1977; NARO, 1981; NARO e SCHERRE, 1999a). Na categoria mais saliente (0 a 2 sílabas), encontramos mais concordância (53,7% de frequência e 0,515 de peso relativo), porque a relação sujeito/SN-verbo é mais óbvia. O índice cai para 43,2% (0,398 de peso relativo) quando temos de três a cinco sílabas, chegando a 21,4% (0,094 de peso relativo) quando temos presença de seis ou mais sílabas entre sujeito/SN e verbo.

Na categoria menos saliente (6 ou mais sílabas), computamos apenas 14 ocorrências, sendo que três trazem a marca de plural nos seus respectivos verbos.

⁵³ No controle desta variável, foram desconsideradas as 350 ocorrências com sujeito nulo e as 160 ocorrências com sujeito/SN posposto ao verbo.

- (106) ah os pai da minha esposa até não::: não **chiaru** comigo... (M0PB)
- (107) eles sempre tipo sempre **acompanharam** (M2PB)
- (108) porque muitas pessoas na hora que eu falo já **falam** “nossa que metido” (M2PB)

Das três ocorrências, duas (107-108) são de informantes do grupo de pessoas com mais escolaridade. Temos dois verbos (106-107) que pertencem à categoria mais saliente da escala da saliência fônica, em que a oposição bem marcada favorece a concordância. Além disso, os três sujeitos apresentam o traço semântico [+ humano]. Enfim, apesar da distância entre o sujeito/SN e o verbo, há outros fatores favorecedores da concordância verbal.

4.4.8 Presença ou ausência do ‘que’ relativo ou complementizador⁵⁴

Como informamos no início desta seção, a variável presença ou ausência do ‘que’ relativo ou complementizador foi selecionada em último lugar, na rodada em que reagrupamos em três níveis as seis categorias da escala da saliência fônica. Na rodada anterior, a variável não havia sido descartada nem selecionada pelo programa Goldvarb-X. Vale destacar que não foi necessário excluir qualquer grupo de fatores para a variável em questão ser selecionada. Os resultados mostrados na tabela 32 legitimam a hipótese e vêm comprovar os de outros estudos (cf. NARO e SCHERRE, 2003a; MONTE, 2007; entre outros).

⁵⁴ No controle desta variável, foram desconsideradas as 350 ocorrências com sujeito nulo e as 160 ocorrências com sujeito/SN posposto ao verbo.

Tabela 32: Frequência e peso relativo de concordância verbal em função da presença ou ausência do 'que' relativo ou complementizador no PB.

Fatores	Frequência	PR
– presença do 'que'	50/125 = 40,0%	0,344
– ausência do 'que'	431/787 = 54,8%	0,526
Total	481/912 = 52,7%	Range 182
<i>Input: 0,464</i>		
Significância: 0,046		

Os índices confirmam que a presença do 'que' mascara, de fato, a relação entre o SN controlador da concordância e o verbo. Com a presença do 'que', as chances de cancelamento da marca de plural são maiores (0,344) do que quando o 'que' está ausente (0,526). A diferença entre os dois fatores em termos de peso relativo é de 182.

Cruzando a variável presença/ausência do 'que' com a variável distância entre o sujeito/SN e o verbo em número de sílabas, podemos verificar melhor a atuação do 'que'. Para esse cruzamento, resolvemos deixar de lado os casos em que não há material fonético entre sujeito/SN e verbo (zero sílaba)⁵⁵.

Tabela 33: frequência de concordância verbal em função do número de sílabas entre o sujeito/SN e o verbo e a presença ou ausência do 'que' no PB.

Fatores	Frequência	
	– presença do 'que'	– ausência do 'que'
– 1 a 2 sílabas	43/103 = 42%	92/146 = 63%
– 3 a 5 sílabas	7/19 = 37%	12/25 = 48%
– 6 ou mais sílabas	0/3 = 0%	3/11 = 27%

Os resultados, conforme se observa na tabela 33, comprovam que a presença do 'que' inibe a marca de plural nos verbos nos três fatores de número de sílabas. Assim como Naro e Scherre (2003a), podemos concluir que as chances de concordância decrescem com o aumento do número de sílabas, mas decrescem ainda mais se uma das sílabas for o 'que' relativo ou complementizador.

⁵⁵ São 605 ocorrências com zero sílaba (sujeito/SN adjacente ao verbo).

4.4.9 Gênero

No nosso estudo com dados do PB, a variável social gênero foi selecionada em quinto lugar pelo programa Goldvarb-X.

Tabela 34: Frequência e peso relativo de concordância verbal conforme o gênero no PB.

Fatores	Frequência	PR
– masculino	323/708 = 45,6%	0,410
– feminino	363/714 = 50,8%	0,589
Total	686/1.422 = 48,2%	Range 179
<i>Input: 0,466</i>		
<i>Significância: 0,042</i>		

A hipótese estabelecida foi confirmada, pois as mulheres da nossa amostra demonstraram mais sensibilidade à variante de prestígio. Os homens empregaram menos as marcas de plural nos verbos (0,410) do que as mulheres (0,589).

Em nosso estudo anterior com outras 20 pessoas residentes numa comunidade da periferia urbana da cidade de São Carlos (MONTE, 2007), a variável gênero também obteve significância estatística. A título de comparação, apresentamos os resultados do nosso trabalho de 2007.

Tabela 35: Frequência e peso relativo de concordância verbal conforme o gênero no PB, no *corpus* sob análise e no nosso *corpus* de 2007.

Fatores	Frequência de concordância			
	<i>Corpus</i> sob análise		<i>Corpus</i> de 2007	
	Frequência	PR	Frequência	PR
– masculino	323/708 = 45,6%	0,410	121/488 = 25%	0,45
– feminino	363/714 = 50,8%	0,589	126/492 = 26%	0,55
Total	686/1.422 = 48,2%		247/980 = 25%	

Em 2007, interpretamos a variável gênero como inoperante, porque os pesos relativos estão muito próximos da neutralidade e as frequências brutas são, praticamente, as mesmas. Hoje não pensamos assim. A variável foi selecionada nos nossos dois estudos e muitas pesquisas sociolinguísticas já comprovaram a atitude mais conservadora das mulheres, quando estão em jogo uma variante de prestígio e uma variante não prestigiada.

Em relação à concordância verbal de 3ª pessoa do plural, várias são as pesquisas com dados do PB que evidenciam a preferência feminina pela variante padrão (cf. SCHERRE e NARO, 1998b; PEREIRA, 2004; GAMEIRO, 2005, 2009; MONTE, 2007; RUBIO, 2008, 2012; entre outras).

4.4.10 Escolaridade

A variável escolaridade foi a que se mostrou mais relevante na análise estatística realizada, sendo a primeira selecionada pelo Programa Goldvarb-X. Os resultados, expostos na tabela 36, são surpreendentes.

Tabela 36: Frequência e peso relativo de concordância verbal de acordo com o grau de escolaridade do informante no PB.

Fatores	Frequência	PR
– não alfabetizados	89/464 = 19,2%	0,148
– ensino fundamental (EJA)	192/480 = 40,0%	0,351
– ensino médio	405/478 = 84,7%	0,910
Total	686/1.422 = 48,2%	Range 762
<i>Input: 0,466</i>		
<i>Significância: 0,042</i>		

Além de ser a variável selecionada em primeiro lugar, apresenta o *range* mais alto da análise (762). De acordo com as nossas expectativas, constatamos que quanto maior a escolaridade, maior é a frequência de concordância. Fica evidente que a influência normativa da escola acaba tendo um papel importante na aquisição da variedade padrão de concordância.

Chama muito a atenção a diferença marcante entre os dois extremos dos fatores. A diferença em termos de frequência é de 65,5 pontos percentuais e, como já salientamos, de 762 em termos de peso relativo.

Os estudantes concluintes do ensino fundamental na EJA revelam um baixo percentual de frequência de concordância na fala informal (40%). Comparando esse grupo com o de jovens do ensino médio, a diferença ainda é bastante grande: 44,7 pontos percentuais em termos de frequência e 559 em termos de peso relativo.

A distância fica um pouco menor, mas ainda assim significativa, entre os estudantes da EJA e os não alfabetizados da nossa amostra: 20,8 pontos percentuais em termos de frequência e 203 em termos de peso relativo.

Os nossos resultados atuais corroboram os da nossa pesquisa anterior (MONTE, 2007)⁵⁶. Para uma comparação efetiva, eliminamos da amostra atual o grupo de jovens do ensino médio e realizamos uma rodada no programa Goldvarb-X, pois no estudo anterior trabalhamos somente com os outros dois grupos de escolaridade. A tabela seguinte traz os resultados das duas amostras da cidade de São Carlos:

Tabela 37: Frequência e peso relativo de concordância verbal de acordo com o grau de escolaridade do informante no PB, no *corpus* sob análise e no nosso *corpus* de 2007.

Fatores	Frequência de concordância			
	Corpus sob análise		Corpus de 2007	
	Frequência	PR	Frequência	PR
– não alfabetizados	89/464 = 19,2%	0,332	94/491 = 19%	0,40
– ensino fundamental (EJA)	192/480 = 40,0%	0,663	153/489 = 31%	0,60
Total	281/944 = 29,8%		247/980 = 25%	

Os índices encontrados nos dois estudos são muito semelhantes. É sabido que os estudantes da EJA (Educação de Jovens e Adultos) têm um tempo reduzido de estudo, visto que a cada seis meses eles concluem uma série. Os informantes dos nossos dois *corpora* eram do curso noturno e todos estavam concluindo a 8ª série. Não controlamos os anos de escolaridade desses estudantes e tampouco o intervalo de tempo em que ficaram longe da escola, já que cada um deles possui uma experiência de vida muito diferente. De qualquer forma, é possível concluir que a escolaridade, mesmo supletiva, influencia o fenômeno variável de concordância verbal. Continuamos achando muito pertinente a reflexão de Bortoni-Ricardo sem, contudo, deixar de pensar que outras questões referentes ao letramento de uma forma mais ampla contribuem nesse processo.

⁵⁶ Vale lembrar que os informantes da pesquisa concluída em 2007 eram de procedência geográfica diversificada, de uma faixa etária de 20 a 40 anos. Das 20 entrevistas realizadas, 9 pessoas eram da região Nordeste e 11 das regiões Sul e Sudeste. Dessas 11, 5 não eram do Estado de São Paulo e apenas 2 eram da cidade de São Carlos.

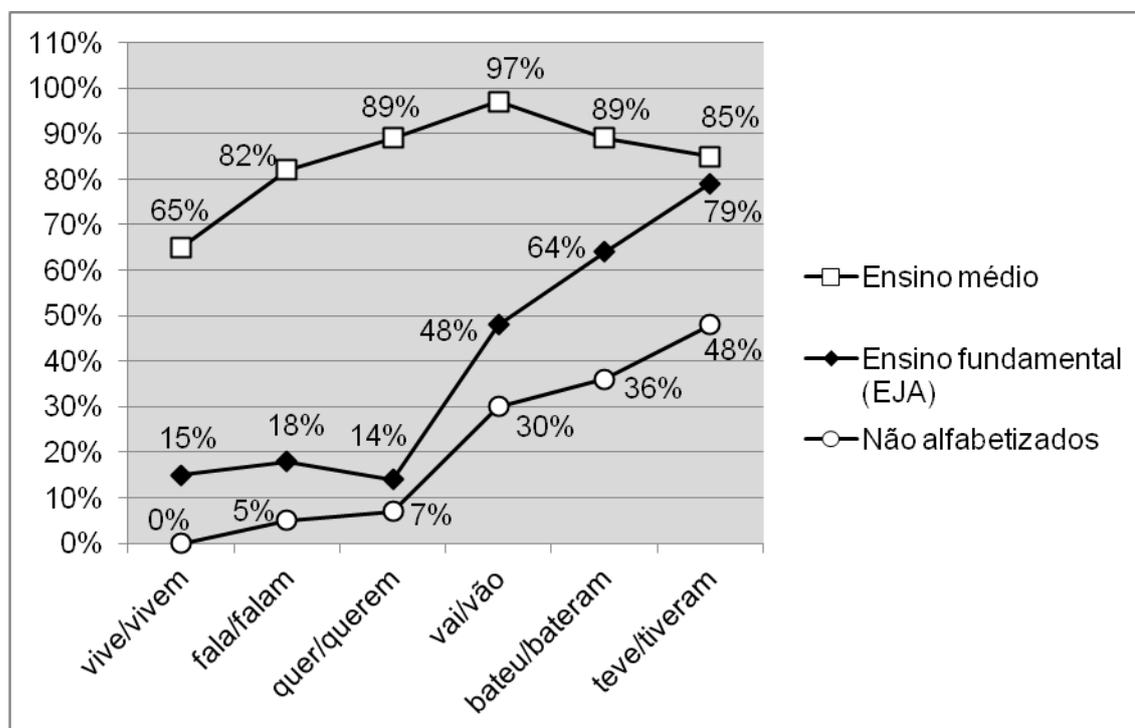
O indivíduo de pouca cultura formal possivelmente só chegará a ter consciência do caráter estigmatizado da concordância não-padrão depois de um período de escolarização e de convívio com o dialeto de classe média urbana. A duração deste período não é possível de se precisar, sem que se proceda a estudos experimentais. **Podemos, porém, prever que esse indivíduo adotará a certa altura de sua formação escolar supletiva as normas de avaliação da classe mais alta, mas tenderá sempre a usar a regra de concordância padrão com menos frequência, posto que ele a assimilou tardiamente.** (BORTONI-RICARDO, 1981, p. 94, grifo nosso).

Assim como Naro e Scherre (1999a), resolvemos analisar o cruzamento das variáveis escolaridade e saliência fônica, a fim de verificar se a nitidez da escala da saliência fônica tem relação com os níveis de escolarização dos informantes. Os resultados se encontram na tabela e no gráfico a seguir:

Tabela 38: Frequência de concordância verbal segundo o grau de escolaridade e o grau de saliência fônica da oposição entre as formas verbais do singular e do plural no PB.

Fatores	Frequência de concordância		
	– não alfabetizados	– ensino fundamental (EJA)	– ensino médio
Nível 1			
– 1a (<i>vive/vivem</i>)	0/22 = 0%	4/27 = 15%	26/40 = 65%
– 1b (<i>fala/falam</i>)	11/239 = 5%	36/205 = 18%	139/169 = 82%
– 1c (<i>quer/querem</i>)	1/15 = 7%	3/21 = 14%	17/19 = 89%
Nível 2			
– 2a (<i>vai/vão</i>)	11/37 = 30%	38/80 = 48%	68/70 = 97%
– 2b (<i>bateu/bateram</i>)	20/56 = 36%	21/33 = 64%	32/36 = 89%
– 2c (<i>teve/tiveram</i>)	46/95 = 48%	90/114 = 79%	123/144 = 85%

Gráfico 11: Frequência de concordância verbal segundo o grau de escolaridade e o grau de saliência fônica da oposição entre as formas verbais do singular e do plural no PB.



Como se pode notar, a escala da saliência fônica é mais nítida para os falantes não alfabetizados e com formação supletiva do que para os jovens do ensino médio. Esses resultados vão ao encontro dos obtidos por Naro e Scherre (1999a) com falantes do Rio de Janeiro. Os pesquisadores verificaram que os dados dos falantes que têm de 1 a 4 anos de escolarização e de 5 a 8 apresentaram uma escala de saliência mais nítida do que os dados dos falantes de 9 a 11 anos de escolarização.

É importante destacar que, com verbos da primeira categoria da oposição não marcada (vive/vivem), os falantes não alfabetizados da nossa amostra apresentam o uso categórico do cancelamento da marca de plural. Na segunda categoria do mesmo nível (fala/falam), encontramos um número maior de ocorrências (239) e um índice ainda muito baixo de frequência de concordância (5%). Para esse grupo de pessoas, o maior salto no índice de concordância ocorre entre os níveis 1 e 2 da hierarquia oposicional. Na última categoria da oposição não marcada (quer/querem), encontramos 7% de frequência de concordância e 30% na primeira categoria da oposição marcada (vai/vão). O salto é de 23 pontos percentuais.

Esse salto é bastante grande também no grupo dos estudantes do ensino fundamental na EJA. O índice de frequência de concordância sobe de 14% para 48%, ao passar da última categoria da oposição não marcada (quer/querem) para a primeira categoria da oposição marcada (vai/vão). A diferença de 34 pontos percentuais é maior do que a diferença encontrada nos não alfabetizados.

Já no grupo de jovens do ensino médio, a escala da saliência fônica é menos nítida, mas ainda assim é atuante. Ao passar do nível 1 para o nível 2, a diferença de 8 pontos percentuais é pequena, comparando com as diferenças dos outros dois grupos. Vemos que o maior índice de concordância ocorre com verbos da primeira categoria da oposição marcada (vai/vão). Nesses contextos, o uso de marcas de plural nos verbos foi quase categórico (97%). Como era esperado, o menor índice de frequência é com verbos da primeira categoria da oposição não marcada (65%). Mas esse índice é alto, perto dos índices encontrados nos não alfabetizados (0%) e nos estudantes do ensino fundamental na EJA (15%).

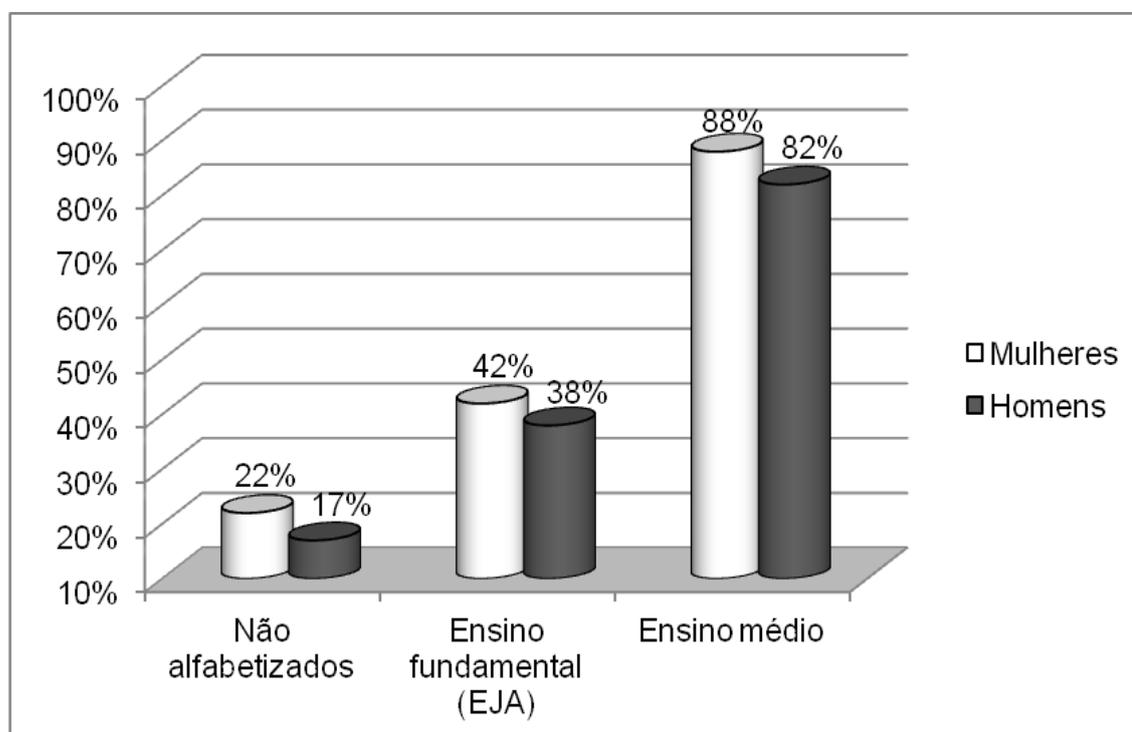
Perante o exposto, podemos concluir que a influência da escolaridade e da escala da saliência fônica na concordância verbal no PB é inquestionável.

Com o objetivo de verificar com maior precisão o papel das mulheres e dos homens no uso da regra de concordância verbal, efetuamos o cruzamento dessas duas variáveis sociais que demonstraram significância estatística.

Tabela 39: Frequência de concordância verbal de acordo com o grau de escolaridade e o gênero do informante no PB.

Fatores	Frequência de concordância	
	– mulheres	– homens
– não alfabetizados	51/234 = 22%	38/230 = 17%
– ensino fundamental (EJA)	101/240 = 42%	91/240 = 38%
– ensino médio	211/240 = 88%	194/238 = 82%

Gráfico 12: Frequência de concordância verbal de acordo com o grau de escolaridade e o gênero do informante no PB.

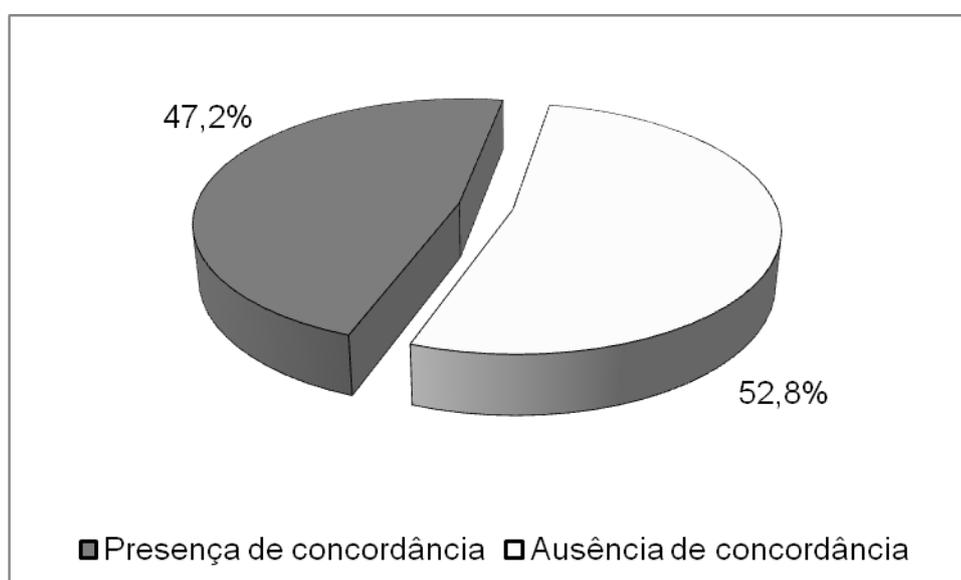


A preferência das mulheres pela variante padrão, prestigiada, comparativamente aos homens, é confirmada nos três níveis de escolaridade por nós analisados. Observamos o aumento na frequência de concordância conforme aumenta o grau de escolaridade dos nossos informantes. Esse movimento ascendente, no entanto, não anula as diferenças estabelecidas pelo gênero, revelando que as duas variáveis agem conjuntamente.

4.5 Análise e discussão dos resultados do PB sem as ocorrências com o verbo 'ser'

Assim como fizemos com os dados do PE, realizamos também uma rodada com os dados do PB sem as ocorrências com o verbo 'ser'. Obtivemos um total de 1.204 dados, com 568 dados (47,2%) apresentando a marca explícita de plural nos verbos e 636 dados (52,8%) com a marca zero de plural nos verbos.

Gráfico 13: Distribuição geral dos dados do PB sem as ocorrências com o verbo 'ser'



Com um *input* 0,448 e significância 0,027, os grupos de fatores selecionados na ordem de relevância foram: escolaridade, saliência fônica, tipo estrutural do sujeito/SN, paralelismo formal no nível oracional, posição do sujeito/SN em relação ao verbo, gênero, distância entre o sujeito/SN e o verbo em número de sílabas e presença ou ausência do 'que' relativo. Nessa rodada, o grupo de fatores traço semântico do sujeito/SN foi descartado pelo *step-down*⁵⁷ e o tipo estrutural do sujeito/SN continuou sendo selecionado pelo *step-up* e excluído pelo *step-down*.

Como os pesos relativos dos fatores são muito semelhantes, comparados aos pesos relativos dos resultados gerais, optamos em apresentar, além da

⁵⁷ Amalgamando os fatores [- humano / + animado] e [- humano / - animado], a variável passa a ser selecionada pelo programa Goldvarb-X. Nesse caso, temos apenas o controle dos traços semânticos [+ humano] e [- humano].

escolaridade, apenas os grupos de fatores que foram selecionados e apresentados com os dados sem o verbo 'ser' do PE, para possíveis comparações.

4.5.1 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural

Dentre as variáveis estruturais, a saliência fônica verbal continua sendo a mais atuante, mesmo sem os dados com o verbo 'ser' (*era/eram; fosse/fossem; será/serão; foi/foram; é/são*).

Tabela 40: Frequência e peso relativo de concordância verbal conforme o grau de saliência fônica no PB, sem as ocorrências com o verbo 'ser'.

Fatores	Frequência	PR
1. nível baixo de saliência fônica (ex.: <i>sai/saem; bate/batem; fala/falam; ganha/ganham</i>)	203/646 = 31,4%	0,236
2. nível médio de saliência fônica (ex.: <i>faz/fazem; tá/tão; bateu/bateram; quer/querem; vai/vão</i>)	199/346 = 57,5%	0,716
3. nível alto de saliência fônica (ex.: <i>quis/quiseram; fez/fizeram; veio/vieram</i>)	166/212 = 78,3%	0,888
Total	568/1.204 = 47,2%	Range 652
<i>Input: 0,448</i>		
<i>Significância: 0,027</i>		

Os índices de frequência e os pesos relativos expostos na tabela 40 revelam que a intensidade dos segmentos fonéticos que realizam a oposição singular/plural exerce um papel importante no condicionamento da concordância.

Os resultados continuam mostrando que as formas verbais no nível baixo de saliência fônica desfavorecem a aplicação da regra de concordância (31,4% de frequência e 0,236 de peso relativo). A aplicação da regra é maior com as formas verbais que apresentam maior diferenciação morfofonológica na oposição singular/plural (nível médio e alto de saliência fônica). Como podemos observar, há um salto bastante grande entre o nível baixo e o nível médio de saliência. Em termos de peso relativo, a diferença entre essas duas categorias é de 480.

4.5.2 Posição do sujeito/SN em relação ao verbo⁵⁸

A tabela a seguir contém os resultados da posição do sujeito/SN em relação ao verbo, sem as ocorrências com o verbo 'ser'.

Tabela 41: Frequência e peso relativo de concordância verbal segundo a posição do sujeito/SN em relação ao verbo no PB, sem as ocorrências com o verbo 'ser'.

Fatores	Frequência	PR
– sujeito/SN anteposto	426/834 = 51,1%	0,553
– sujeito/SN posposto	10/56 = 17,9%	0,040
Total	436/890 = 49,0%	Range 513
<i>Input: 0,448</i>		
<i>Significância: 0,027</i>		

Há, de fato, uma tendência a não interpretar o SN posposto como sujeito (pelo menos não prototípico). Quando o SN vem posposto ao verbo, a frequência de aplicação da regra de concordância verbal cai de 51,1% para 17,9% (uma diferença de 33,2 pontos percentuais). O *range* continua alto (513), comparativamente aos resultados gerais desse grupo de fatores (545).

Os exemplos (109-114) ilustram a ausência de concordância verbal com SN posposto, sem o verbo 'ser'.

- (109) **ia** minhas prima... ia eu minha irmã meu irmão... (F0PB)
- (110) **vai** dois ônibus da Paraty (M0PB)
- (111) e **tá** os treze vivo ainda cê acredita?... (os treze = irmãos) (F1PB)
- (112) acho que **vai** as duas pra lá (as duas = as filhas) (M1PB)
- (113) **vai** muitas pessoas ruins (F2PB)
- (114) **foi** uns dez quinze (só ali)... (uns dez quinze = alunos) (M2PB)

4.5.3 Traço semântico do sujeito/SN

Como informamos inicialmente, nessa rodada, sem as ocorrências com o verbo 'ser', o grupo de fatores traço semântico do sujeito/SN passa a ser selecionado pelo programa Goldvarb-X quando amalgamamos os fatores

⁵⁸ No controle desta variável, foram desconsideradas as 314 ocorrências com sujeito nulo.

[- humano / + animado] e [- humano / - animado]. Dessa forma, não estamos controlando a animacidade, mas apenas os traços semânticos [+ humano] e [- humano].

Tabela 42: Frequência e peso relativo de concordância verbal em função do traço semântico do sujeito/SN no PB, sem as ocorrências com o verbo 'ser'.

Fatores	Frequência	PR
- SN [+ humano]	557/1.135 = 49,1%	0,516
- SN [- humano]	11/69 = 15,9%	0,253
Total	568/1.204 = 47,2%	<i>Range</i> 263
<i>Input: 0,449</i>		
<i>Significância: 0,024</i>		

Os índices de frequência e os pesos relativos mostrados na tabela 42 validam, novamente, a hipótese formulada: o sujeito/SN [+ humano] tende a favorecer a aplicação da regra de concordância (49,1% de frequência e 0,516 de peso relativo), enquanto o sujeito/SN [- humano] tende a inibir a sua aplicação (15,9% de frequência e 0,253 de peso relativo).

4.5.4 Escolaridade

Mesmo eliminando as ocorrências com o verbo 'ser' da rodada, a variável escolaridade continua sendo a primeira selecionada pelo programa Goldvarb-X, com o *range* mais alto (752).

Tabela 43: Frequência e peso relativo de concordância verbal de acordo com o grau de escolaridade do informante no PB, sem as ocorrências com o verbo 'ser'.

Fatores	Frequência	PR
- não alfabetizados	76/397 = 19,1%	0,159
- ensino fundamental (EJA)	156/416 = 37,5%	0,354
- ensino médio	336/391 = 85,9%	0,911
Total	568/1.204 = 47,2%	<i>Range</i> 752
<i>Input: 0,448</i>		
<i>Significância: 0,027</i>		

Constatam-se praticamente os mesmos índices de frequência de concordância e os mesmos pesos relativos, comparativamente aos resultados gerais dessa variável (cf. tabela 36).

Ainda é muito baixa a frequência de concordância verbal (19,1%) na fala das pessoas não alfabetizadas (0,159 de peso relativo). O índice de frequência aumenta para 37,5% (0,354 de peso relativo) na fala das pessoas com formação supletiva (EJA). Já os falantes que frequentaram a escola por mais tempo, o grupo do ensino médio, realizam mais a concordância verbal (85,9% de frequência e 0,911 de peso relativo).

4.6 Comparação dos resultados do PB e do PE

Nesta seção, retomamos os grupos de fatores que foram selecionados igualmente nas amostras do PB e do PE, realizando uma comparação entre eles. O quadro 1, a seguir, mostra a ordem de seleção e o *range* dos grupos de fatores nas duas amostras.

Quadro 1: A ordem de seleção e o *range* dos grupos de fatores no PB e no PE.

Grupos de fatores		São Carlos (PB)		Évora (PE)	
		Ordem de seleção	Range	Ordem de seleção	Range
Linguísticos	Saliência fônica	2º	716	5º	425
	Tipo estrutural do sujeito/SN	3º	537	3º	608
	Paralelismo formal oracional	4º	543	Não selecionado	–
	Posição do sujeito/SN	6º	545	1º	344
	Traço semântico do sujeito/SN	7º	248	2º	260
	Verbo 'ser' / Outros verbos	8º	157	4º	339
	Distância entre S e V (sílabas)	9º	421	Não selecionado	–
	Presença/ausência do 'que'	10º	182	Não selecionado	–
Sociais	Escolaridade	1º	762	Não selecionado	–
	Gênero	5º	179	6º	140

4.6.1 Grau de saliência fônica da oposição entre as formas verbais do singular e do plural

A saliência fônica foi a variável estrutural mais relevante na análise com os dados do PB e foi selecionada em quinto lugar na análise com os dados do PE. Observa-se, na tabela seguinte, que o *range* é mais alto (716) na variedade brasileira.

Tabela 44: Frequência e peso relativo de concordância verbal conforme o grau de saliência fônica da oposição entre as formas verbais do singular e do plural, na amostra do PB e do PE.

Fatores	Frequência de concordância verbal			
	São Carlos (PB)		Évora (PE)	
	Frequência	PR	Frequência	PR
Nível 1				
– 1a (<i>vive/vivem</i>)	30/89 = 33,7%	0,142	119/130 = 91,5%	0,299
– 1b (<i>fala/falam</i>)	186/613 = 30,3%	0,224	536/581 = 92,3%	0,380
– 1c (<i>quer/querem</i>)	21/55 = 38,2%	0,213	86/88 = 97,7%	0,427
Nível 2				
– 2a (<i>vai/vão</i>)	117/187 = 62,6%	0,723	147/154 = 95,5%	0,477
– 2b (<i>bateu/bateram</i>)	73/125 = 58,4%	0,805	81/86 = 94,2%	0,637
– 2c (<i>teve/tiveram</i>)	259/353 = 73,4%	0,858	371/401 = 92,5%	0,724
Total	686/1.422 = 48,2%	Range 716	1.340/1.440 = 93,1%	Range 425

A escala da saliência fônica na sua primeira dimensão (oposição não marcada, desfavorecendo a concordância, vs. oposição marcada, favorecendo a concordância) é mais nítida para os dados do português brasileiro. O salto do nível 1 para o nível 2 é muito significativo nessa variedade. Já no PE, a passagem do nível 1 para o nível 2 não é tão demarcada.

Além disso, os valores mostrados na tabela 44 evidenciam claramente a diferença quantitativa na variação da concordância verbal de 3ª pessoa do plural no PB e no PE. Destacamos essas diferenças nas três categorias do primeiro nível. Há uma diferença de 57,8 pontos percentuais na primeira categoria da oposição não marcada (*vive/vivem*) entre as duas variedades. Seguindo a escala, a diferença na segunda categoria (*fala/falam*) entre o PB e o PE é de 62 pontos percentuais. E na última categoria da oposição não marcada (*quer/querem*), a diferença entre as duas variedades é de 59,5 pontos percentuais.

O menor peso relativo nas variedades brasileira e europeia é justamente na primeira categoria do nível 1, onde a única marca de plural é uma nasalização da vogal final inacentuada.

Encontramos duas possíveis interpretações do efeito da saliência fônica na concordância verbal no português brasileiro. Para Lemle e Naro (1977, p.49), a atuação da mudança em direção a um sistema sem concordância verbal no PB foi fundamentalmente fonológica. Teve início primeiramente nas categorias menos salientes e foi se difundindo nas categorias cada vez mais salientes. Os autores argumentam que a perda da nasalidade final inacentuada é um fenômeno bem mais

amplo na língua portuguesa, possível de se encontrar no Brasil e em Portugal, em variedades antigas e modernas.

De fato, quando vamos às obras da dialetologia portuguesa, é comum encontrarmos exemplos que ilustram o fenômeno da desnasalização na sílaba final da palavra. Em um trabalho recente de Simão (2011) sobre o falar de Marvão⁵⁹, vemos que com várias palavras o fenômeno ocorre (barragem, homem, linguagem, ontem, coragem, forragem, garagem, mensagem, moagem, nuvens, ordens, passagem, Portagem, vargem).

Guy (2005), por sua vez, dá uma interpretação contrária à de Lemle e Naro sobre a saliência fônica e a concordância verbal no PB.

[...] a escala de saliência funciona como uma escala de facilidade de aquisição. Na aquisição de uma regra por uma pessoa que não tem essa regra, a escala de saliência atuaria no sentido de que ela adquiriria primeiro os casos mais óbvios, mais salientes, e os mais difíceis de aprender seriam adquiridos só depois. Assim, esses dados podem ser interpretados como evidência de uma história em que havia uma comunidade de pessoas que não tinham essa regra de concordância, e que a estavam adquirindo através do contato com falantes nativos, que mostravam a regra nas suas produções orais. Este é o caso que existiu na história das línguas crioulas, durante os processos de pidginização e descrioulização. (GUY, 2005, p. 26).

Lemle e Naro (1977) veem um processo de perda de concordância em curso. O PB estaria caminhando em direção a um sistema sem concordância previsto pela deriva secular. Já Guy (2005) pensa em aquisição, ou seja, teríamos saído de um sistema sem concordância (ou com pouca), caminhando para um sistema com concordância – a ideia de descrioulização.

4.6.2 Tipo estrutural do sujeito/SN

O grupo de fatores tipo estrutural do sujeito/SN foi selecionado em terceiro lugar nas duas amostras. Nesse grupo de fatores, a variedade europeia apresenta o *range* mais alto (608).

⁵⁹ Cidade portuguesa que pertence ao Distrito de Portalegre, localizada no Alto Alentejo.

Tabela 45: Frequência e peso relativo de concordância verbal de acordo com o tipo estrutural do sujeito/SN, na amostra do PB e do PE.

Fatores	Frequência de concordância verbal			
	São Carlos (PB)		Évora (PE)	
	Frequência	PR	Frequência	PR
– Pronome <i>vocês</i>	14/19 = 73,7%	0,583	8/8 = 100%	–
– Pronome pessoal <i>eles/elas</i>	311/503 = 61,8%	0,565	168/169 = 99,4%	0,805
– Numeral (núcleo ou adjunto)	37/125 = 29,6%	0,660	76/85 = 89,4%	0,664
– Sujeito nulo	162/350 = 46,3%	0,500	506/526 = 96,2%	0,527
– Pronome demonstrativo	3/10 = 30,0%	0,123	35/37 = 94,6%	0,462
– Quantificador indefinido	39/83 = 47,0%	0,306	72/78 = 92,3%	0,436
– SN pleno simples	95/284 = 33,5%	0,400	390/436 = 89,4%	0,370
– SN pleno nu	14/22 = 63,6%	0,540	71/84 = 84,5%	0,294
– SN pleno composto	11/26 = 42,3%	0,309	14/17 = 82,4%	0,197
Total	686/1.422 = 48,2%	Range 537	1.340/1.440 = 93,1%	Range 608

Uma semelhança que encontramos nas duas variedades é o pronome ‘vocês’, o pronome pessoal ‘eles/elas’ e o numeral (núcleo ou adjunto) favorecendo a concordância verbal. Mas as diferenças em termos de frequência, em todos os fatores, são bem marcadas. Como se vê, há o uso praticamente categórico da regra de concordância com o pronome ‘ele/elas’ (99,4%), na amostra do PE. O mesmo não ocorre no PB, mas ainda assim é uma das categorias com o maior índice de concordância verbal (61,8%). Destacamos as diferenças nas categorias com pronome demonstrativo e com SN pleno nu. Na variedade brasileira, o pronome demonstrativo aparece desfavorecendo a concordância verbal (0,123) e, na variedade europeia, se encontra em uma faixa intermediária (0,462). O SN pleno nu preserva mais a marca de plural no verbo no *corpus* do PB (0,540) do que no *corpus* do PE (0,294).

Uma das diferenças que é sempre apontada entre o português brasileiro e o português europeu é o uso mais frequente de pronomes na função de sujeito. O número de ocorrências nas duas amostras estudadas revela essa tendência. São 503 dados de pronome ‘eles/elas’ na amostra do PB, que correspondem a 35,4%, contra 169 dados na amostra do PE, que correspondem a 11,7%. Já o uso de sujeito nulo é maior na variedade europeia. São 526 dados na amostra do PE contra 350 dados na amostra do PB, que correspondem a 36,5% e 24,6%, respectivamente.

Pode-se observar na tabela que o índice de frequência de concordância no PB é maior quando o pronome ‘eles/elas’ está presente (61,8%) do que quando o sujeito é nulo (46,3%). Segundo Naro e Scherre (2003c, p.291), “este fato reforça a

dissociação entre o fenômeno de diminuição no uso de desinências verbais distintas e o aumento no uso de pronomes sujeito”.

Chegamos a pensar que os informantes com mais escolaridade da nossa amostra do PB utilizassem com menos frequência o sujeito pronominal ‘eles/elas’. Não foi o que encontramos.

Tabela 46: Frequência de uso do sujeito pronominal ‘eles/elas’ e de sujeito nulo segundo o grau de escolaridade na amostra do PB.

Fatores	– não alfabetizados	– ensino fundamental (EJA)	– ensino médio
– sujeito ‘eles/elas’	155/503 = 30,8%	160/503 = 31,8%	188/503 = 37,4%
– sujeito nulo	110/350 = 31,4%	143/350 = 40,9%	97/350 = 27,7%

Nota-se um aumento discreto do uso do pronome ‘eles/elas’ na fala dos informantes com escolaridade mais elevada. E o uso mais elevado de sujeito nulo está no nível intermediário de escolaridade (40,9%). Do nível intermediário para o nível mais alto de escolaridade, a percentagem de uso de sujeito nulo cai para 27,7%.

Já vimos que quanto mais escolaridade, mais se aplica a concordância verbal. No entanto, os nossos dados do PB também indicam que esse aumento na concordância não está correlacionado com a diminuição de realização do sujeito pronominal (cf. NARO e SCHERRE, 2003c; PAREDES SILVA, 2008).

4.6.3 Posição do sujeito/SN em relação ao verbo

Para o PE, a posição do sujeito em relação ao verbo foi o primeiro grupo de fatores selecionado e, para o PB, foi o sexto a ser selecionado pelo programa Goldvarb-X.

Tabela 47: Frequência e peso relativo de concordância verbal em função da posição do sujeito/SN em relação ao verbo, na amostra do PB e do PE.

Fatores	Frequência de concordância verbal			
	São Carlos (PB)		Évora (PE)	
	Frequência	PR	Frequência	PR
– sujeito/SN anteposto	481/912 = 52,7%	0,613	680/715 = 95,1%	0,581
– sujeito/SN posposto	43/160 = 26,9%	0,068	154/199 = 77,4%	0,237
Total	524/1.072 = 48,9%	<i>Range 545</i>	834/914 = 91,2%	<i>Range 344</i>

Como podemos notar, a contribuição dessa variável para a explicação da variação é bastante relevante nas duas amostras. No PB, foi selecionada em sexto lugar, como já informamos, mas o *range* é mais alto (545), comparado ao *range* dessa variável na amostra do PE (344).

Brandão e Vieira (2012) analisaram 1.515 dados de construções de 3ª pessoa do plural de um *corpus* da região metropolitana de Lisboa e encontraram apenas dezessete dados sem a marca de concordância verbal (1,1%). Da análise qualitativa que fizeram desses dados, as autoras concluem o seguinte:

As construções com sujeito posposto constituem, na realidade, um dos poucos contextos estruturais, amplamente debatido na literatura, em que as três variedades igualmente registram casos de cancelamento da marca de número plural. (BRANDÃO e VIEIRA, 2012, p. 1056)⁶⁰.

4.6.4 Traço semântico do sujeito/SN

Na tabela a seguir, podemos observar a relevância estatística do grupo de fatores traço semântico do sujeito/SN na amostra do PB e do PE.

Tabela 48: Frequência e peso relativo de concordância verbal em função do traço semântico do sujeito/SN, na amostra do PB e do PE.

Fatores	Frequência de concordância verbal			
	São Carlos (PB)		Évora (PE)	
	Frequência	PR	Frequência	PR
– SN [+ humano / + animado]	656/1.284 = 51,1%	0,522	1.082/1.130 = 95,8%	0,555
– SN [– humano / + animado]	0/24 = 0,0%	–	28/31 = 90,3%	0,438
– SN [– humano / – animado]	30/114 = 26,3%	0,274	230/279 = 82,4%	0,295
Total	686/1.422 = 48,2%	Range 248	1.340/1.440 = 93,1%	Range 260

Esse grupo de fatores foi o sétimo selecionado para a amostra do PB, com um *range* não muito alto (248). No PE, foi selecionado como estatisticamente relevante em segundo lugar, mas apresentou o segundo menor *range* (260).

O que vemos nos resultados da tabela 48 é o traço semântico [+ humano] favorecendo a realização da concordância e os sujeitos/SNs inanimados favorecendo a ausência de concordância.

⁶⁰ As três variedades mencionadas se referem ao Português de Portugal, do Brasil e de São Tomé.

No fator intermediário, SN [– humano / + animado], a hipótese foi confirmada apenas no PE. No PB, ocorreu o uso categórico de não concordância. Há um indício de que o traço [humano] é mais relevante do que o traço [animado]. Mas temos de relativizar essa análise, levando em consideração os poucos dados na categoria e o fato de não ter ocorrido esse SN na fala dos informantes do nível mais alto de escolaridade (cf. subseção 4.4.5).

4.6.5 Tipo de verbo: verbo ‘ser’ versus outros verbos

O grupo de fatores tipo de verbo (verbo ‘ser’ versus outros verbos) foi selecionado nas duas amostras, mas com comportamento diferente. No PB, ficou em oitavo lugar com o menor *range* (157) e, no PE, foi selecionado em quarto lugar com o quarto maior *range* (339).

Tabela 49: Frequência e peso relativo de concordância verbal segundo o tipo de verbo (verbo ‘ser’ e outros verbos), na amostra do PB e do PE.

Fatores	Frequência de concordância verbal			
	São Carlos (PB)		Évora (PE)	
	Frequência	PR	Frequência	PR
– verbo ‘ser’	118/218 = 54,1%	0,632	291/342 = 85,1%	0,247
– outros verbos	568/1.204 = 47,2%	0,475	1.049/1.098 = 95,5%	0,586
Total	686/1.422 = 48,2%	<i>Range</i> 157	1.340/1.440 = 93,1%	<i>Range</i> 339

Nossas expectativas se confirmaram apenas para a amostra do PE, com o verbo ‘ser’ desfavorecendo a concordância verbal (0,247 de peso relativo). Já na nossa amostra do PB, sentenças com o verbo ‘ser’ favorecem a aplicação da regra de concordância (0,632 de peso relativo). A diferença dos pesos relativos nessa categoria chama a atenção.

Nas gramáticas normativas há sempre uma seção na parte de concordância verbal destinada ao verbo ‘ser’. Rocha Lima (1998, p.404), por exemplo, intitula a seção de “concordância especial do verbo ‘ser’”. Como já mencionamos, o ‘ser’ é um verbo que aparece nas mais diversas construções da língua portuguesa: em sentenças apresentacionais, atributivas e equativas. Segundo Castilho (2010, p. 398-399), é um verbo que desfruta de grande frequência de uso, seja como verbo funcional, seja como verbo auxiliar. Vai dizer ainda que o uso apresentativo de ‘ser’ representa uma sobrevivência de seu uso como verbo pleno.

De fato, é um verbo especial e tem de ser controlado separadamente. Em relação à concordância verbal de 3ª pessoa do plural com o verbo 'ser', os nossos resultados evidenciam um comportamento diferente entre o PB e o PE, que necessita de um estudo quantitativo e qualitativo mais detalhado.

4.6.6 Gênero

A variável gênero foi selecionada em quinto lugar na amostra do PB e em último lugar na amostra do PE. Os resultados da tabela 50 mostram que as mulheres utilizam mais a forma padrão do que os homens, nas duas amostras estudadas.

Tabela 50: Frequência e peso relativo de concordância verbal segundo a variável gênero, na amostra do PB e do PE.

Fatores	Frequência de concordância verbal			
	São Carlos (PB)		Évora (PE)	
	Frequência	PR	Frequência	PR
– masculino	323/708 = 45,6%	0,410	663/720 = 92,1%	0,430
– feminino	363/714 = 50,8%	0,589	677/720 = 94,0%	0,570
Total	686/1.422 = 48,2%	Range 179	1.340/1.440 = 93,1%	Range 140

Na variedade europeia, a ordem de seleção e o menor *range* (140) revelam que os grupos de fatores estruturais são mais atuantes na definição da variação. Na variedade brasileira, embora tenha sido selecionado em quinto lugar, apresenta o segundo menor *range* (179). Vemos que, em termos de peso relativo, os índices quase se igualam nos dois *corpora*.

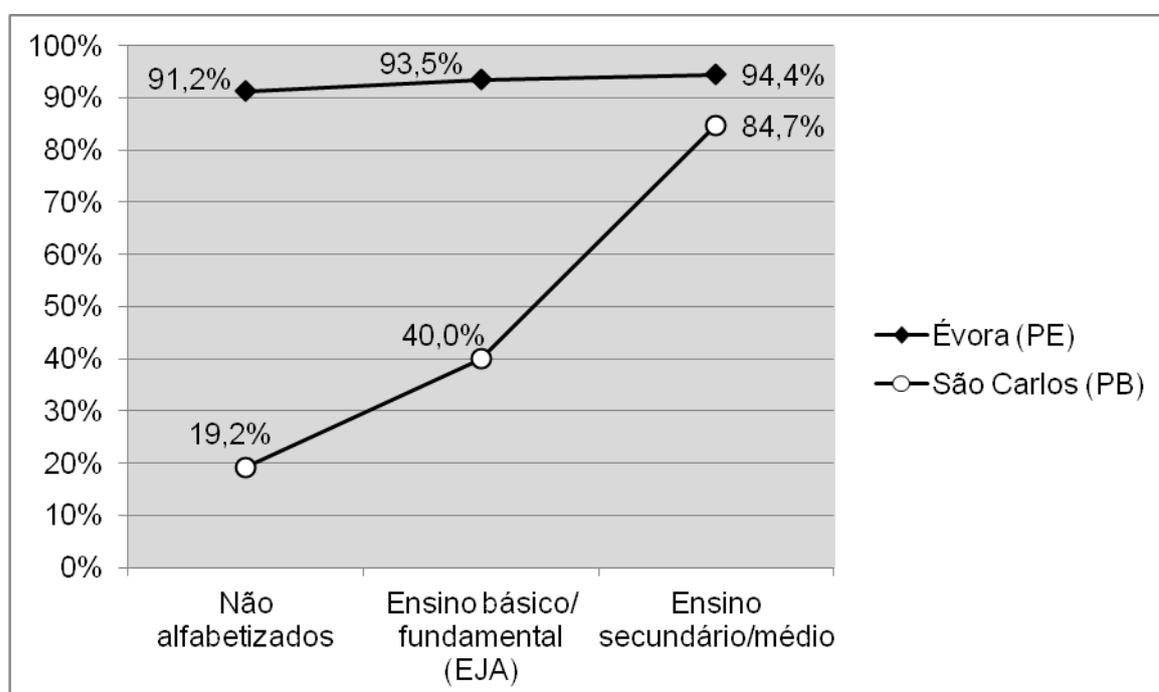
A nossa investigação comprova, juntamente com outras, a importância de se controlar a variável gênero nos estudos sociolinguísticos. Contudo, de acordo com Paiva (2004):

A questão que se coloca para o sociolinguista é a de explicar os padrões regulares depreendidos em diferentes pesquisas e a natureza das possíveis diferenças linguísticas entre homens e mulheres. É necessário cuidado para não tomar como fatos indicações que só podem ser interpretadas no plano simbólico. (PAIVA, 2004, p. 39).

4.6.7 Escolaridade

A variável escolaridade foi selecionada como a mais relevante estatisticamente para a amostra do PB, com o *range* mais alto (762). Já para a amostra do PE, a escolaridade não obteve significância estatística. Resolvemos mostrar os resultados dessa variável nas duas amostras, porque as diferenças encontradas “saltam aos olhos”:

Gráfico 14: Frequência de concordância verbal segundo a escolaridade, na amostra do PB e do PE.



A diferença em termos de frequência entre os não alfabetizados das duas amostras é de 72 pontos percentuais. Entre os informantes da faixa intermediária, a diferença é um pouco menor (53,5 pontos percentuais), mas ainda muito significativa. O distanciamento já não é tão acentuado na faixa de informantes com maior escolarização (9,7 pontos percentuais de diferença).

Em termos quantitativos, é possível dizer que os informantes de Évora apresentam uma regra de aplicação da concordância verbal praticamente semicategórica (cf. LABOV, 2003, p. 243).

No nosso *corpus* do Brasil, a regra de concordância verbal na fala informal das pessoas é, de fato, uma regra variável. Nem mesmo na fala dos jovens com mais escolaridade da nossa amostra a regra se configura como semicategórica.

Essas diferenças, além de surpreendentes, nos fazem refletir sobre a complexa realidade sociolinguística do Brasil. Por que, em Portugal, não há uma polarização como temos no Brasil?

De acordo com Lucchesi, Baxter e Silva (2009, p. 348), essa polarização sociolinguística ocorre nas diferentes variedades do português brasileiro, constituindo um contínuo. O contínuo exemplificado pelo autores tem num extremo as comunidades rurais afro-brasileiras do interior do Estado da Bahia, onde o nível de aplicação da regra de concordância é da ordem de 16%, e no outro extremo a norma urbana culta do Rio de Janeiro, com uma frequência de 94% de aplicação da regra (cf. GRACIOSA, 1991 *apud* LUCCHESI, BAXTER e SILVA, 2009).

O cotejo dos dois extremos do *continuum* constitui prova contundente da polarização sociolinguística do Brasil: enquanto o uso da regra é quase categórico entre falantes urbanos com elevado grau de escolaridade, o nível de variação é elevadíssimo nas comunidades rurais afro-brasileiras [...] (LUCCHESI, BAXTER e SILVA, 2009, p.348).

Não encontramos essa polarização sociolinguística apenas entre diferentes variedades do PB. Ela está presente também dentro de uma mesma comunidade de fala, como bem indicam os nossos dados da cidade de São Carlos. Sendo assim, é possível pensar em um contínuo de nível de instrução (cf. BRANDÃO e VIEIRA, 2012). Com os índices encontrados na nossa amostra do PB, o contínuo nível de instrução fica assim representado:

– marcas		+ marcas
-----	-----	-----
Não alfabetizados	Fundamental (EJA)	Fundamental
19,2%	40,0%	84,7%
	–	–
		Médio
		Superior

O contínuo de nível de instrução, proposto por Brandão e Vieira (2012), vem enriquecer a ideia de Bortoni-Ricardo (2004), que propõe três contínuos (contínuo de urbanização; contínuo de oralidade–letramento; contínuo de monitoração estilística) para entender a variação no português do Brasil.

+ rural — — — — — + urbano
 + oralidade — — — — — + letramento
 – monitorado — — — — — + monitorado

Não realizamos um estudo sistemático dos nossos dados com base nos quatro contínuos acima mencionados. Entretanto, pensando nos informantes não alfabetizados da nossa amostra do PB, por exemplo, é possível perceber suas características gerais nesses quatro contínuos. São analfabetos, por isso o evento é de oralidade. Hoje vivem na cidade, mas tiveram uma forte relação com a zona rural. Nesses casos, Bortoni-Ricardo (2004) vai dizer que são *rurbanos*⁶¹, pela forte influência rural na cultura e na língua. Quanto ao contínuo de monitoração estilística, notamos um estilo menos monitorado, devido às características das entrevistas sociolinguísticas informais.

Os informantes eborenses sem qualquer escolaridade também têm essas características. São senhoras e senhores que não sabem ler nem escrever, tiveram uma forte ligação com a zona rural (nos campos alentejanos), hoje vivem na cidade e as entrevistas sociolinguísticas foram conduzidas da mesma maneira.

Por que, então, essa diferença gritante nas frequências de aplicação da regra de concordância verbal no PB e no PE?

Não conseguiremos responder essa pergunta se recusarmos o desenvolvimento histórico da sociedade em que a língua é usada.

Diante dos nossos resultados, fica muito difícil negar a influência decisiva do massivo contato entre línguas que marca os primeiros séculos da história sociolinguística do Brasil (cf. LUCCHESI, 2003, 2006, 2008; LUCCHESI e BAXTER, 2009).

Nesse cenário histórico de polarização sociolinguística, os processos de variação e mudança induzidos pelo contato entre línguas teriam ocorrido na formação histórica da norma popular brasileira e só indiretamente, mediante o que os sociolinguistas têm chamado de *contato dialetal*, afetaram o desenvolvimento histórico da norma culta brasileira, podendo explicar também por que são tantas as diferenças entre o português brasileiro e o português europeu, mesmo quando a observação se atém aos falantes ditos cultos. (LUCCHESI, 2008, p.368).

⁶¹ Segundo a pesquisadora, a terminologia é da antropologia social.

5 CONSIDERAÇÕES FINAIS

Por meio de um estudo empírico e sincrônico, nos moldes da Sociolinguística Variacionista Laboviana, analisamos a variação na concordância verbal de terceira pessoa do plural na fala de 18 pessoas residentes na cidade de São Carlos, localizada no interior do Estado de São Paulo/Brasil, e na fala de 18 pessoas residentes na cidade de Évora, situada no Alentejo, sul de Portugal. Em cada comunidade de fala, os informantes foram selecionados em função do gênero (9 mulheres e 9 homens) e da escolaridade (6 não alfabetizados, 6 concluintes do ensino fundamental/básico na EJA (Educação de Jovens e Adultos) e 6 concluintes do ensino médio/secundário).

Os nossos resultados com dados do português europeu (PE) revelam uma significativa diferença quantitativa em relação aos nossos resultados com dados do português brasileiro (PB). Na amostra do PB, encontramos 48,2% (686/1.422) de presença de concordância verbal. Já na amostra do PE, 93,1% (1.340/1.440) dos dados trazem a marca explícita de plural nos verbos. De acordo com a tipologia de regras apresentada por Labov (2003, p. 242), podemos afirmar que no português do Brasil a regra é efetivamente variável. No português de Portugal, a regra parece ter *status* semicategórico.

Controlamos, neste estudo, oito variáveis linguísticas e duas variáveis sociais (gênero e escolaridade), com o objetivo de verificar quais estariam condicionando ou restringindo a variação na concordância verbal de 3ª pessoa do plural nas duas comunidades estudadas.

Na análise dos resultados do PB, todas as variáveis linguísticas e sociais revelaram-se estatisticamente significativas no condicionamento da regra variável de concordância verbal, a saber: (1) saliência fônica verbal; (2) tipo estrutural do sujeito/SN; (3) paralelismo formal no nível oracional; (4) posição do sujeito/SN em relação ao verbo; (5) traço semântico do sujeito/SN; (6) tipo de verbo (verbo 'ser' *versus* outros verbos); (7) distância entre o sujeito/SN e o verbo em número de sílabas; (8) presença ou ausência do 'que' relativo ou complementizador; (9) escolaridade e (10) gênero.

No PE, não foram todas as variáveis linguísticas e sociais selecionadas. Das oito variáveis linguísticas controladas, cinco se mostraram relevantes no

condicionamento da variação: (1) posição do sujeito/SN em relação ao verbo; (2) traço semântico do sujeito/SN; (3) tipo estrutural do sujeito/SN; (4) tipo de verbo (verbo 'ser' *versus* outros verbos) e (5) saliência fônica verbal. Dos dois grupos de fatores sociais, o gênero foi selecionado.

Pelo controle da variável saliência fônica verbal, verificamos que as chances de o verbo apresentar marca explícita de plural são maiores quando a diferença entre a forma singular e plural é mais perceptível, ou seja, mais saliente. A escala da saliência se mostrou mais nítida na amostra do PB. O princípio da saliência fônica tem sido amplamente confirmado nos estudos desenvolvidos com dados do português do Brasil. Foi a variável estrutural mais significativa na nossa análise com dados do PB. No PE, foi selecionada em quinto lugar e não são todos os estudos que revelam a sua atuação na variação da concordância verbal (cf. MONGUILHOTT, 2009; RUBIO, 2012).

Alterações fonético-fonológicas são comuns nas línguas naturais. Segundo Coelho (1967 [1880], p. 104), “a transformação da linguagem em virtude da alteração fonética é um fenômeno de base fisiológica”. As obras da dialetologia portuguesa sempre trazem exemplos de variação de natureza fonético-fonológica. Florêncio (2005) apresentou um estudo com base nas recolhas dialetológicas realizadas por Leite de Vasconcellos no Alentejo (Portugal), entre 1883 e 1936. Em relação às terminações verbais, Florêncio (2005, p. 65-66) comenta que “a terminação *-am* da terceira pessoa do plural do presente do indicativo dos verbos da 1ª conjugação e do pretérito imperfeito das 2ª e 3ª conjugações realiza-se [õ] (*andôm, andavôm, andõ, cômprõ*)” e que “a terminação nasal *-em* da 3ª pessoa do plural do presente do indicativo dos verbos das 2ª e 3ª conjugações, bem como do presente e do imperfeito do conjuntivo dos verbos da 1ª conjugação, realiza-se [ã] ou [êj] (*vêndê, bébê, bébêm, fazêm*)”. É importante dizer que o trabalho de Florêncio (2005) não trata de qualquer dado como ausência de concordância. No entanto, essas particularidades no nível fonético-fonológico podem interferir na realização da concordância verbal, e o controle da saliência fônica verbal, no nosso trabalho, mostra essa tendência.

As variáveis posição do sujeito/SN em relação ao verbo e traço semântico do sujeito/SN são muito atuantes na variação da concordância verbal, tanto na variedade brasileira quanto na variedade europeia. Em relação ao PE, sabemos que há divergências quanto à seleção das variáveis, isto é, nem sempre as mesmas

variáveis são significativas nos diferentes *corpora*. Entretanto, a posição do sujeito/SN em relação ao verbo e o traço semântico do sujeito/SN foram as únicas variáveis selecionadas nos cinco trabalhos com dados do PE que utilizaram o programa Goldvarb⁶². Vejamos:

Quadro 2: Variáveis selecionadas em diferentes *corpora* do PE.

Trabalhos sobre concordância verbal em diferentes <i>corpora</i> do PE	Variáveis selecionadas
Oito textos portugueses medievais (NARO e SCHERRE, 2007)	<ul style="list-style-type: none"> – Saliência fônica verbal – Posição do sujeito/SN em relação ao verbo – Traço semântico do sujeito/SN
Lisboa (MONGUILHOTT, 2009)	<ul style="list-style-type: none"> – Traço semântico do sujeito/SN – Posição do sujeito em relação ao verbo – Tipo de verbo – Idade/escolaridade
Funchal – Ilha da Madeira (BAZENGA, 2011)	<ul style="list-style-type: none"> – Posição do sujeito/SN em relação ao verbo – Traço semântico do sujeito/SN – Saliência fônica verbal – Escolaridade
CRPC – Portugal (RUBIO, 2012)	<ul style="list-style-type: none"> – Posição do sujeito/SN em relação ao verbo – Traço semântico do sujeito/SN – Tipo estrutural do sujeito/SN
Évora (resultados do presente trabalho)	<ul style="list-style-type: none"> – Posição do sujeito/SN em relação ao verbo – Traço semântico do sujeito/SN – Tipo estrutural do sujeito/SN – Tipo de verbo (verbo ‘ser’ <i>versus</i> outros verbos) – Saliência fônica verbal – Gênero

Até mesmo os trabalhos que não utilizaram o programa Goldvarb, mas que fizeram uma análise qualitativa dos dados de ausência de concordância encontrados no PE destacaram a posposição do sujeito/SN como um contexto estrutural favorável ao cancelamento da marca de plural nos verbos. Mas, de acordo com Brandão e Vieira (2012, p. 1056), “há que se questionar o estatuto dessas construções quanto à identidade do sujeito”. No PB, desde o trabalho pioneiro de Eunice Pontes (1986), vários estudos vêm questionando o estatuto do SN posposto

⁶² Varejão (2006) utilizou o programa Goldvarb, mas não apresentou resultados em termos de pesos relativos. Os demais trabalhos (cf. CARRILHO, 2003; MOTA e VIEIRA, 2008; GANDRA, 2009; ALMEIDA, 2010; VIEIRA, 2012; BRANDÃO e VIEIRA, 2012) realizaram uma análise qualitativa das ocorrências de ausência de concordância verbal.

ao verbo. Lembramos que nos casos dos SNs introduzidos pelos verbos apresentacionais, Castilho (2010, p. 288), na *Nova Gramática do Português Brasileiro*, propõe classificá-los como absolutivo. Castilho (2010) e Bagno (2011) argumentam que esses sintagmas nominais não têm propriedades sintáticas de sujeito nem de objeto. Acreditamos que essa interpretação seja válida para o português de Portugal também. Se considerarmos que a concordância verbal é entendida como um processo ligado à identificação do sujeito, a ausência de marcas de concordância com tais SNs pospostos seria previsível, dado o seu estatuto sintático. Esse contexto não diferenciaria, então, o PB e o PE.

Resolvemos analisar o verbo 'ser' separadamente dos outros verbos. A hipótese levantada foi confirmada apenas para a amostra do PE, com o verbo 'ser' inibindo mais a concordância verbal do que os outros verbos. Na amostra do PB, o verbo 'ser' aparece favorecendo a aplicação da regra de concordância. Diante dos resultados, concluímos que é legítimo tratar o verbo 'ser' separadamente e que as diferenças de resultados entre as duas variedades podem ter a ver com os tipos de construção com 'ser' que ocorreram no PB e no PE. Num trabalho futuro, pretendemos realizar um estudo qualitativo e quantitativo mais detalhado do verbo 'ser', que permita apreender melhor o seu comportamento.

A variável gênero se mostrou relevante na amostra brasileira e na amostra lusitana, com as mulheres utilizando mais a forma padrão de concordância do que os homens. Esse resultado já era esperado no Brasil, pois vários estudos sociolinguísticos revelam que as mulheres usam mais a forma de prestígio, quando temos uma variante socialmente prestigiada e outra desprestigiada. Com dados do PE, é o primeiro estudo sociolinguístico sobre a variação da concordância verbal de 3ª pessoa do plural que demonstra relevância do gênero. Para avançarmos na discussão, temos de entender o papel das mulheres são-carlenses e das mulheres eborenses na sociedade em que estão inseridas. Além disso, são necessários outros estudos desta natureza em terras lusitanas que atestem a relevância do fator gênero.

Os grupos de fatores distância entre o sujeito/SN e o verbo em termos do número de sílabas e presença/ausência do 'que' relativo ou complementizador foram selecionados apenas para a amostra do PB. Esse fato indica que no PB a presença da concordância está mais atrelada a condições de proximidade. Qualquer elemento que rompe a proximidade (a adjacência) entre o sujeito/SN e o verbo fragiliza a

identificação de um vínculo que já não é muito perceptível, principalmente para o falante com nula ou pouca escolaridade. Isso também se percebe na atuação do grupo de fatores paralelismo formal no nível oracional, que não foi selecionado para o PE, mas foi selecionado em quarto lugar para o PB. Os nossos resultados mostram que há uma regra mais consistente do ponto de vista estrutural no PE para a concordância, que vai se pautar em uma percepção mais nítida da função sujeito ou da ideia de que a marca de plural no verbo é um modo de marcar a relação desse verbo com o sujeito. Vemos que aspectos mais ligados ao processamento linear da relação sujeito-verbo (paralelismo, adjacência/ruptura) atuam na variação da concordância verbal no PB, mas não afetam a concordância verbal no PE.

A variável que mais se mostrou relevante na amostra do PB foi a escolaridade. Já na amostra lusitana, a variável não obteve significância estatística. Na amostra brasileira, vemos nitidamente a crescente frequência de concordância à medida que aumenta a escolarização. Na amostra do PE, as diferenças nos índices de frequência são irrelevantes do ponto de vista estatístico.

	– marcas		+ marcas
PB	-----		
	Não alfabetizados 19,2%	Fundamental (EJA) 40,0%	Ensino Médio 84,7%

	– marcas		+ marcas
PE	-----		
	Não alfabetizados 91,2%	Ensino Básico (EJA) 93,5%	Ensino Secundário 94,4%

Os nossos resultados da atuação da escolaridade apontam para a ideia de que efetivamente o funcionamento da concordância no PB é diferente do PE. No Brasil, ao contrário do que ocorre em Portugal, há uma polarização sociolinguística entre diferentes variedades e dentro de uma mesma comunidade. A fala de uma pessoa sem qualquer escolaridade traz muito o cancelamento da marca de plural nos verbos. A escola, por sua vez, acaba tendo um papel fundamental e decisivo na modificação do comportamento linguístico do falante. Sabemos que a concordância verbal é um fenômeno variável no PB que não está imune à estigmatização e que

gozam de um prestígio social maior os falantes que a utilizam. Consequentemente, é um tópico gramatical que não escapa à atenção normativa da escola.

Considerando os resultados obtidos com a escolaridade no PB, desejamos, num futuro trabalho de pesquisa, estudar as relações entre variação, ensino, aprendizagem e norma linguística. Nossa proposta é investigar como a escola lida com o ensino da concordância verbal e como está ocorrendo esta aprendizagem por meio dos anos de escolarização, levando em consideração outras questões referentes ao letramento de uma forma mais ampla.

Resta dizer que esta pesquisa não tem a pretensão de responder se a variação na concordância verbal de terceira pessoa do plural no Brasil é fruto da deriva secular inerente na língua trazida de Portugal (cf. NARO e SCHERRE, 2007) ou do processo de transmissão linguística irregular (cf. LUCCHESI, 2003, 2006, 2008; LUCCHESI e BAXTER, 2009). Mas acreditamos que os estudos descritivo-comparativos como este podem contribuir nas discussões sobre as origens do português do Brasil. Um fato inquestionável é que o massivo contato entre línguas na história de nossa formação sociolinguística pode explicar as notáveis diferenças entre o português brasileiro e o português europeu. Temos de continuar o garimpo e, de acordo com Galves (2008) e Petter (2009), outros *corpora*, inclusive de variedades africanas, devem entrar na nossa agenda de pesquisa. Segundo Naro e Scherre (2007, p. 85), “são desconhecidas a verdadeira extensão e a intensidade da variação na concordância em terras lusitanas”. O nosso estudo vem colaborar com esse garimpo e é mais uma peça importante desse complexo e instigante quebra-cabeça. E o debate continua...

6 REFERÊNCIAS

- ALMEIDA, E. M. **Uso e norma**: variação da concordância verbal em redações escolares. 2010. Tese (Doutorado) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.
- BAGNO, M. **Não é errado falar assim!** Em defesa do português brasileiro. São Paulo: Parábola Editorial, 2009.
- BAGNO, M. **Gramática pedagógica do português brasileiro**. São Paulo: Parábola Editorial, 2011.
- BAXTER, A. N. A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a crioulação prévia: um exemplo do estado da Bahia. In: D'ANDRADE, E.; KIHM, A. (Orgs.). **Actas do colóquio sobre crioulos de base lexical portuguesa**. Lisboa: Edições Colibri, 1992. p. 7-36.
- BAZENGA, A. M. Concordância verbal e variantes de 3ª pessoa do plural em PE: resultados preliminares de um estudo sociolinguístico com base numa amostra de português falado no Funchal. In: SILVA, A. S.; TORRES, A.; GONÇALVES, M. (Orgs.). **Línguas pluricêntricas**: variação linguística e dimensões sociocognitivas. Braga: Publicações da Faculdade de Filosofia, Universidade Católica Portuguesa, 2011. p. 302-318.
- BELINE, R. A variação linguística. In: FIORIN, J. L. (Org.). **Introdução à linguística**. I. Objetos teóricos. 4. ed. São Paulo: Contexto, 2005. p. 121-140.
- BERLINCK, R. A. **A ordem V SN no português do Brasil**: sincronia e diacronia. 1988. Dissertação (Mestrado) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1988.
- BORTONI-RICARDO, S. M. A concordância verbal em português: um estudo de sua significação social. In: COUTO, H. H. (Ed.). **Ensaio de linguística aplicada ao português**. Brasília: Thesaurus, 1981. p. 79-101.
- BORTONI-RICARDO, S. M. **Educação em língua materna**: a sociolinguística na sala de aula. São Paulo: Parábola Editorial, 2004.
- BORTONI-RICARDO, S. M. A concordância verbal em português: um estudo de sua significação social. In: VOTRE, S.; RONCARATI, C. (Orgs.). **Anthony Julius Naro e a linguística no Brasil**: uma homenagem acadêmica. Rio de Janeiro: 7Letras, 2008. p. 363-380.
- BORTONI-RICARDO, S. M. **Do campo para a cidade**: estudo sociolinguístico de migração e redes sociais. Tradução Stella Maris Bortoni-Ricardo, Maria do Rosário Rocha Caxangá. São Paulo: Parábola Editorial, 2011.

BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. **Alfa: Revista de Linguística**, São Paulo, v. 56, n. 3, p. 1035-1064, 2012.

BRIGHT, W. As dimensões da sociolinguística. In: FONSECA, M. S. V.; NEVES, M. F. (Orgs.). **Sociolinguística**. Rio de Janeiro: Eldorado, 1974. p. 17-23. (Coleção Enfoque, 3).

CÂMARA JUNIOR, J. M. Línguas europeias de ultramar: o português do Brasil. In: _____. **Dispersos de J. Mattoso Câmara Jr.** Rio de Janeiro: Lucerna, 2004 [1972]. p. 113-126.

CARRILHO, E. Ainda a “unidade e diversidade da língua portuguesa”: a sintaxe. In: CASTRO, I.; DUARTE, I. (Eds.). **Razões e emoção**: miscelânea de estudos em homenagem a Maria Helena Mira Mateus. Lisboa: Imprensa Nacional – Casa da Moeda, 2003. v. 2, p. 19-41.

CASTILHO, A. T. **Nova gramática do português brasileiro**. São Paulo: Contexto, 2010.

CHAMBERS, J. K.; TRUDGIL, P. **La dialectologia**. Traducción Carmen Morán Gonzalez. Madrid: Visor Libros, 1994 [1980].

COELHO, A. Os dialectos românicos ou neo-latinos na África, Ásia e América. In: ACADEMIA INTERNACIONAL DA CULTURA PORTUGUESA. **Estudos linguísticos**: crioulos. Reedição de artigos publicados no Boletim da Sociedade de Geografia de Lisboa. Lisboa, 1967 [1881]. p. 1-234.

CUNHA, C.; CINTRA, L. F. L. **Nova gramática do português contemporâneo**. 3.ed. Rio de Janeiro: Nova Fronteira, 2001.

DECAT, M. B. N. Concordância verbal, topicalização e posposição do sujeito. In: SIMÕES, A. M.; REIS, C. A. (Orgs.). **Cadernos de linguística e teoria da literatura**. Belo Horizonte: UFMG, 1983. ano V, n. 9.

FERREIRA, V. **Carta ao futuro**: ensaio. Lisboa: Quetzal, 2010 [1958].

FISCHER, J. L. Influências sociais na escolha de variantes linguísticas. In: FONSECA, M. S. V.; NEVES, M. F. (Orgs.). **Sociolinguística**. Rio de Janeiro: Eldorado, 1974. p. 87-98. (Coleção Enfoque, 3).

FLORÊNCIO, M. **Dialecto alentejano**: contributos para o seu estudo. 2. ed. Lisboa: Edições Colibri, 2005.

GALVES, C. O enfraquecimento da concordância no português brasileiro. In: ROBERTS, I.; KATO, M. A. (Orgs.). **Português brasileiro**: uma viagem diacrônica: homenagem a Fernando Tarallo. Campinas: Editora da Unicamp, 1993. p. 387-408.

GALVES, C. O papel das línguas africanas na formação do português brasileiro: (mais) pistas para uma nova agenda de pesquisa. **Gragoatá**, Niterói, n. 24, p. 145-164, 2008.

GAMEIRO, M. B. **A concordância verbal na língua falada da região central do estado de São Paulo**. 2005. Dissertação (Mestrado) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2005.

GAMEIRO, M. B. **A variação da concordância verbal na terceira pessoa do plural em redações escolares do ensino fundamental e médio**: uma avaliação de fatores linguísticos e sociais. 2009. Tese (Doutorado) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2009.

GONÇALVES, S. C. L. **O português falado na região de São José do Rio Preto**: constituição de um banco de dados anotado para o seu estudo. Projeto de pesquisa apresentado à FAPESP. 2003. Disponível em: <<http://www.iboruna.ibilce.unesp.br/histórico/projeto>>.

GONÇALVES, S. C. L. **O português falado na região de São José do Rio Preto**: constituição de um banco de dados anotado para o seu estudo. Relatório científico parcial III apresentado à FAPESP. 2007. Disponível em: <<http://www.iboruna.ibilce.unesp.br/histórico/relatorio3>>.

GANDRA, A. S. A concordância verbal no português europeu rural. In: OLIVEIRA, K.; CUNHA e SOUZA, H. F.; GOMES, L. (Orgs.). **Novos tons de Rosa... para Rosa Virgínia Mattos e Silva**. Salvador: EDUFBA, 2009. p. 142-161.

GUY, G. R. A identidade linguística da comunidade de fala: paralelismo interdialeto nos padrões de variação linguística. **Organon**, Porto Alegre, v. 14, n. 28-29, p.17-32, 2000.

GUY, G. R. A questão da crioulização no português do Brasil. In: ZILLES, A. M. S. (Org.). **Estudos de variação linguística no Brasil e no Cone Sul**. Porto Alegre: Editora da UFRGS, 2005. p. 15-38.

GUY, G. R.; ZILLES, A. **Sociolinguística quantitativa**: instrumental de análise. São Paulo: Parábola Editorial, 2007.

ILARI, R.; BASSO, R. **O português da gente**: a língua que estudamos a língua que falamos. São Paulo: Contexto, 2006.

LABOV, W. **Sociolinguistic patterns**. Philadelphia: University of Pennsylvania Press, 1972.

LABOV, W. **Where does the linguistic variable stop?** A response to Beatriz Lavandera. Working papers in sociolinguistics, number 44. Austin: Southwest Educational Development Laboratory, 1978.

LABOV, W. **Principles of linguistic change**. Malden: Blackwell Publishers, 1994. v. 1: Internal factors.

LABOV, W. **Principios del cambio lingüístico**. Madrid: Gredos, 1996. v. 1: Factores internos.

LABOV, W. **Principles of linguistic change**. Malden: Blackwell Publishers, 2001. v. 2: Social factors.

LABOV, W. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Eds.). **Sociolinguistics: the essential readings**. Malden: Blackwell Publishing, 2003. p. 234-250.

LABOV, W. **Padrões sociolinguísticos**. Tradução Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008 [1972].

LAVANDERA, B. R. Where does the linguistic variable stop? **Language in Society**, Cambridge, v. 7, p. 171-182, 1978.

LAVANDERA, B. R. Los limites de la variable sociolingüística. In: _____. **Variación y significado**. Buenos Aires: Libreria Hachette, 1984 [1978]. p. 37-46.

LEMLE, M.; NARO, A. J. **Competências básicas do português**. Relatório final da pesquisa apresentado às instituições patrocinadoras Fundação Movimento Brasileiro de Alfabetização (Mobral) e Fundação Ford. Rio de Janeiro: Mobral, Fundação Ford, 1977.

LUCCHESI, D. O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In: RONCARATI, C.; ABRAÇADO, J. (Orgs.). **Português brasileiro: contato linguístico, heterogeneidade e história**. Rio de Janeiro: 7Letras, 2003. p. 272-284.

LUCCHESI, D. **Sistema, mudança e linguagem: um percurso na história da linguística moderna**. São Paulo: Parábola Editorial, 2004.

LUCCHESI, D. Parâmetros sociolinguísticos do português brasileiro. **Revista da ABRALIN**, Brasília, v. 5, n. 1-2, p. 83-112, 2006.

LUCCHESI, D. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto. In: RONCARATI, C.; ABRAÇADO, J. (Orgs.). **Português brasileiro II: contato linguístico, heterogeneidade e história**. Niterói: EdUFF, 2008. p. 366-390.

LUCCHESI, D. A metodologia. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). **O português afro-brasileiro**. Salvador: EDUFBA, 2009. p. 155-164.

LUCCHESI, D.; BAXTER, A. A transmissão linguística irregular. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). **O português afro-brasileiro**. Salvador: EDUFBA, 2009. p. 101-124.

LUCCHESI, D.; BAXTER, A.; SILVA, J. A. A. A concordância verbal. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Orgs.). **O português afro-brasileiro**. Salvador: EDUFBA, 2009. p. 331-372.

MATEUS, M. H. M. et al. **Gramática da língua portuguesa**. 6. ed. Lisboa: Editorial Caminho, 2003.

MATEUS, M. H. M. Maria Helena Mira Mateus: a dama da linguística portuguesa. Entrevistadores: Paulo Bearzoti Filho e Carlos Alberto Faraco. **Discutindo Língua Portuguesa**, São Paulo, v. 1, n. 3, p. 8-11, 2006.

MATTOS e SILVA, R. V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola Editorial, 2004.

MATTOS e SILVA, R. V. **Caminhos da linguística histórica: ouvir o inaudível**. São Paulo: Parábola Editorial, 2008.

MONGUILHOTT, I. O. S. **Variação na concordância verbal de terceira pessoa do plural na fala dos florianopolitanos**. 2001. Dissertação (Mestrado) – Universidade Federal de Santa Catarina, Florianópolis, 2001.

MONGUILHOTT, I. O. S. **Estudo sincrônico e diacrônico da concordância verbal de terceira pessoa do plural no PB e no PE**. 2009. Tese (Doutorado) – Universidade Federal de Santa Catarina, Florianópolis, 2009.

MONTE, A. **Concordância verbal e variação: uma fotografia sociolinguística da cidade de São Carlos**. 2007. Dissertação (Mestrado) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2007.

MONTE, A. A variação da concordância verbal na modalidade falada e escrita da língua portuguesa do Brasil. In: SILVA, A. S.; TORRES, A.; GONÇALVES, M. (Orgs.). **Línguas pluricêntricas: variação linguística e dimensões sociocognitivas**. Braga: Publicações da Faculdade de Filosofia, Universidade Católica Portuguesa, 2011. p. 502-508.

MOTA, M. A.; VIEIRA, S. R. Contrastando variedades do português brasileiro e europeu: padrões de concordância sujeito-verbo. In: GONÇALVES, C. A.; ALMEIDA, M. L. L. (Orgs.). **Língua portuguesa: identidade, difusão e variabilidade**. Rio de Janeiro: AILP – UFRJ, 2008. p. 111-137.

NARO, A. J. The social and structural dimensions of a syntactic change. **Language**, Baltimore, v. 57, n. 1, p. 63-98, 1981.

NARO, A. J.; LEMLE, M. Syntactic diffusion. **Ciência e Cultura**, São Paulo, v. 29, n. 3, p. 259-268, 1977.

NARO, A. J.; SCHERRE, M. M. P. Variação e mudança linguística: fluxos e contrafluxos na comunidade de fala. **Cadernos de Estudos Linguísticos**, Campinas, n. 20, p. 9-16, 1991.

NARO, A. J.; SCHERRE, M. M. P. Influência de variáveis escalares na concordância verbal. **A cor das letras**. Feira de Santana: Universidade Estadual de Feira de Santana, 1999a. p. 17-34.

NARO, A. J.; SCHERRE, M. M. P. Sobre o efeito do princípio da saliência na concordância verbal na fala moderna, na escrita antiga e na escrita moderna. In: MOURA, D. (Org.). **Os múltiplos usos da língua**. Maceió: EDUFAL, 1999b. p. 26-37.

NARO, A. J.; SCHERRE, M. M. P. A hierarquização do controle da concordância no português moderno e medieval: o caso de estruturas de sujeito composto. In: GROßE, S.; ZIMMERMANN, K. (Eds.). **O português brasileiro**: pesquisas e projetos. Frankfurt am Main: TFM, 2000. v. 17, p. 166-188.

NARO, A. J.; SCHERRE, M. M. P. A relação verbo/sujeito: o efeito máscara do *que* relativo. In: HORA, D.; COLLISCHONN, G. (Orgs.). **Teoria linguística**: fonologia e outros temas. João Pessoa: Editora Universitária, 2003a. p. 383-401.

NARO, A. J.; SCHERRE, M. M. P. Estabilidade e mudança linguística em tempo real: a concordância de número. In: PAIVA, M. C.; DUARTE, M. E. L. (Orgs.). **Mudança linguística em tempo real**. Rio de Janeiro: Contra Capa Livraria, 2003b. p. 47-62.

NARO, A. J.; SCHERRE, M. M. P. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In: RONCARATI, C.; ABRAÇADO, J. (Orgs.). **Português brasileiro**: contato linguístico, heterogeneidade e história. Rio de Janeiro: 7Letras, 2003c. p. 285-302.

NARO, A. J.; SCHERRE, M. M. P. **Origens do português brasileiro**. São Paulo: Parábola Editorial, 2007.

NARO, A. J.; SCHERRE, M. M. P. Fluxos e contrafluxos – movimentos sociolinguísticos da comunidade de fala brasileira. In: MOLLICA, M. C. M. (Org.). **Usos da linguagem e sua relação com a mente humana**. Rio de Janeiro: Tempo Brasileiro, 2010. p. 79-90.

NASCENTES, A. **O idioma nacional**. 3. ed. Rio de Janeiro: Livraria Acadêmica, 1960.

NOLL, V. **O português brasileiro**: formação e contrastes. Tradução do alemão por Mario Eduardo Viaro. São Paulo: Globo, 2008.

OLIVEIRA, N. C. **A concordância verbal na região noroeste do estado de São Paulo**. 2010. Dissertação (Mestrado) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 2010.

OLIVEIRA e SILVA, G. M.; PAIVA, M. C. A. Visão de conjunto das variáveis sociais. In: OLIVEIRA e SILVA, G. M.; SCHERRE, M. M. P. (Orgs.). **Padrões sociolinguísticos**: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro: Departamento de Linguística e Filologia, UFRJ, 1996. cap. 16, p. 335-378.

PAIVA, M. C. A variável gênero/sexo. In: MOLLICA, M. C.; BRAGA, M. L. (Orgs.). **Introdução à sociolinguística: o tratamento da variação**. 2. ed. São Paulo: Contexto, 2004. p. 33-42.

PAIVA, M. C.; DUARTE, M. E. L. Quarenta anos depois: a herança de um programa na sociolinguística brasileira [Posfácio]. In: WEINREICH, U.; LABOV, W.; HERZOG, M. I. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução Marcos Bagno. São Paulo: Parábola Editorial, 2006. p. 131-151.

PAREDES SILVA, V. L. Ainda sobre sujeitos pronominais e reflexão verbal: revisitando Naro. In: VOTRE, S.; RONCARATI, C. (Orgs.). **Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica**. Rio de Janeiro: 7Letras, 2008. p. 381-392.

PEREIRA, D. C. **Concordância verbal na língua falada nas trilhas das bandeiras paulistas**. 2004. Dissertação (Mestrado) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

PESSOA, F. **Mensagem**. Lisboa: Assírio & Alvim, 2010 [1934].

PETTER, M. M. T. O continuum afro-brasileiro do português. In: GALVES, C.; GARMES, H.; RIBEIRO, F. R. (Orgs.). **África-Brasil: caminhos da língua portuguesa**. Campinas: Editora da Unicamp, 2009. p. 159-173.

PINTZUK, S. **Varbrul programs**. 1988. Mimeografado.

PONTES, E. S. L. **Sujeito: da sintaxe ao discurso**. São Paulo: Ática, 1986.

RIZZOLI, A. **A escravidão em São Carlos: inventário analítico**. São Carlos: Universidade Federal de São Carlos, Núcleo de Estudos Afro-Brasileiros, 1995.

ROCHA LIMA, C. H. **Gramática normativa da língua portuguesa**. 36. ed. Rio de Janeiro: José Olympio, 1998.

RODRIGUES, A. C. S. **A concordância verbal no português popular em São Paulo**. 1987. Tese (Doutorado) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo.

RODRIGUES, A. C. S. Concordância verbal e estratégias de pronominalização. Estudos linguísticos. In: SEMINÁRIOS DO GEL, 18., 1989. **Anais...** Lorena: GEL / Prefeitura Municipal de Lorena, 1989. p. 546-554.

RODRIGUES, A. C. S. Concordância verbal e saliência social no português popular de São Paulo. In: GARTNER, E.; HUNDT, C.; SCHOMBERG (Eds.). **Estudos de sociolinguística brasileira e portuguesa**. Frankfurt am Main: TFM, 2000. p. 41-61. (Biblioteca luso-brasileira, 15).

RONCARATI, C. Prestígio e preconceito linguísticos. **Cadernos de Letras da UFF**, Niterói, n. 36, p. 45-56, 2008.

RUBIO, C. F. **A concordância verbal na língua falada na região noroeste do estado de São Paulo**. 2008. Dissertação (Mestrado) – Universidade Estadual Paulista, São José do Rio Preto, 2008.

RUBIO, C. F. **Padrões de concordância verbal e de alternância pronominal no português brasileiro e europeu: estudo sociolinguístico comparativo**. 2012. Tese (Doutorado) – Universidade Estadual Paulista, São José do Rio Preto, 2012.

SANKOFF, D. Variable rules. In: AMMON, U.; DITTMAR, N.; MATTHEIER, K. J. (Eds.). **Sociolinguistics: an international handbook of the science of language and society**. Berlin: Walter de Gruyter, 1988. p. 984-998.

SÃO CARLOS. Prefeitura Municipal. **Plano Diretor do município de São Carlos: Lei nº 13.691 de 25 de novembro de 2005**. São Carlos: Prefeitura Municipal, 2005. 1CD-ROM.

SCHERRE, M. M. P. Paralelismo linguístico. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 7, n. 2, p. 29-59, 1998.

SCHERRE, M. M. P. **Doa-se lindos filhotes de poodle: variação linguística, mídia e preconceito**. São Paulo: Parábola Editorial, 2005.

SCHERRE, M. M. P.; NARO, A. J. Duas dimensões do paralelismo formal na concordância verbal no português popular do Brasil. **D.E.L.T.A.**, São Paulo, v. 9, n. 1, p. 1-14, 1993.

SCHERRE, M. M. P.; NARO, A. J. Restrições sintáticas e semânticas no controle da concordância verbal em português. **Fórum linguístico**, Florianópolis, v. 1, n. 1, p. 45-71, 1998a.

SCHERRE, M. M. P.; NARO, A. J. Sobre a concordância de número no português falado do Brasil. In: RUFFINO, G. (Org.). **Dialettologia, geolinguística, sociolinguística**. In: CONGRESSO INTERNAZIONALE DI LINGUISTICA E FILOLOGIA ROMANZA, 21., 1998, Centro di Studi Filologici e Linguistici Siciliani, Università di Palermo. Tübingen: Max Niemeyer Verlag, 1998b. v. 5, p. 509-523.

SCHERRE, M. M. P.; NARO, A. J. A hierarquização do controle da concordância no português moderno e medieval: o caso de estruturas de sujeito simples. In: GROßE, S.; ZIMMERMANN, K. (Eds.). **O português brasileiro: pesquisas e projetos**. Frankfurt am Main: TFM, 2000. v. 17, p. 135-165.

SCHERRE, M. M. P.; NARO, A. J. Passado e presente na concordância de número em português: evidências do português europeu moderno. In: MASSINI-CAGLIARI, G.; MURAKAWA, C. A. A.; BERLINCK, R. A.; GUEDES, M. (Orgs.). **Estudos de linguística histórica do português**. Araraquara: Laboratório Editorial da FCL/UNESP; São Paulo: Cultura Acadêmica Editora, 2005. p. 31-71.

SCHERRE, M. M. P.; NARO, A. J.; CARDOSO, C. R. O papel do tipo de verbo na concordância verbal no português brasileiro. **D.E.L.T.A.**, São Paulo, v. 23, n. esp., p. 283-317, 2007.

SILVA NETO, S. **Introdução ao estudo da língua portuguesa no Brasil**. 3. ed. Rio de Janeiro: Presença, 1976.

SIMÃO, T. S. B. **O falar de Marvão**: pronúncia, vocabulário, alcunhas, ditados e provérbios populares. Lisboa: Edições Colibri, 2011.

SIMPLÍCIO, M. D. V. M. **Evolução e morfologia do espaço urbano de Évora**. 1997. Tese (Doutorado) – Universidade de Évora, Évora, 1997.

TAGLIAMONTE, S. A. **Analysing sociolinguistic variation**. Cambridge: Cambridge University Press, 2006.

TAGLIAMONTE, S. A. Quantitative analysis. In: BAYLEY, R.; LUCAS, C. (Eds.). **Sociolinguistic variation**: theories, methods and applications. Cambridge: Cambridge University Press, 2007. p. 190-214.

TARALLO, F. **A pesquisa sociolinguística**. 7. ed. São Paulo: Ática, 2002.

TRASK, R. L. **Dicionário de linguagem e linguística**. Tradução Rodolfo Ilari. 2. ed. São Paulo: Contexto, 2006.

TRUZZI, O. **Café e indústria**: São Carlos: 1850-1950. 3. ed. São Carlos: EdUFSCar, 2007.

VAREJÃO, F. O. A. **Variação em estruturas de concordância verbal e em estratégias de relativização no português europeu popular**. 2006. Tese (Doutorado) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

VIEIRA, S. R. **Concordância verbal**: variação em dialetos populares do norte fluminense. 1995. Dissertação (Mestrado) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1995.

VIEIRA, S. R. Estatuto da regra variável e o fenômeno da concordância verbal em variedades de português. In: CONGRESO INTERNACIONAL DE LA ALFAL, 16., 2011, Alcalá de Henares. **Actas...** Alcalá de Henares: Servicio de Publicaciones de la Universidad de Alcalá, 2012. p. 4376-4383. Disponível em: <<http://alfal2011.mundoalfal.org/>>.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. Empirical Foundations for a theory of language change. In: LEHMANN, W. P.; MALKIEL, Y. (Eds.). **Directions for historical linguistics**. Austin-London: University of Texas Press, 1968. p. 95-195.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. **Fundamentos empíricos para uma teoria da mudança linguística**. Tradução Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].

ZILLES, A. M. S. A posposição do sujeito ao verbo no português falado no Rio Grande do Sul. **Letras de Hoje**, Porto Alegre, v. 35, n. 1, p. 75-96, 2000.