

UNIVERSIDADE ESTADUAL DO SUDOESTE DA BAHIA (UESB)
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGÜÍSTICA (PPGLIN)

GILBERTO ALMEIDA MEIRA

**ESTUDO COMPARATIVO ENTRE AS NORMAS POPULAR E CULTA DO
PORTUGUÊS DE VITÓRIA DA CONQUISTA: CONCORDÂNCIA NOMINAL DE
NÚMERO**

VITÓRIA DA CONQUISTA – BA

2015

GILBERTO ALMEIDA MEIRA

**ESTUDO COMPARATIVO ENTRE AS NORMAS POPULAR E CULTA DO
PORTUGUÊS DE VITÓRIA DA CONQUISTA: CONCORDÂNCIA NOMINAL DE
NÚMERO**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística (PPGLIN) da Universidade Estadual do Sudoeste da Bahia (UESB), como requisito parcial e obrigatório para obtenção do título de Mestre em Linguística.

Área de Concentração: Linguística

Linha de Pesquisa: Descrição e análise de Línguas Naturais

Orientador: Prof. Dr. Jorge Augusto Alves da Silva

Coorientadora: Prof^ª Dra. Valéria Viana Sousa

VITÓRIA DA CONQUISTA – BA

2015

Meira, Gilberto Almeida.

M451e Estudo comparativo entre as normas popular e culta do português de Vitória da Conquista: concordância nominal de número / Gilberto Almeida Meira, 2015.
131f.: il. ; algumas col.

Orientador (a): Jorge Augusto Alves da Silva.

Coorientadora: Profa. Dra. Valéria Sousa Viana.

Dissertação (mestrado) – Universidade Estadual do Sudoeste da Bahia, Programa de Pós-graduação em Linguística, Vitória da Conquista, 2015.

Referências: f. 124-128.

1. Concordância nominal de número. 2. Variação linguística (Linguagem culta e popular) – Vitória da Conquista(BA). I. Silva, Jorge Augusto Alves da. II. Universidade Estadual do Sudoeste da Bahia, Programa de Pós-Graduação em Linguística. III.T.

CDD: 469.5

Catálogo na fonte: Elinei Carvalho Santana – CRB 5/1026
UESB – Campus Vitória da Conquista – BA

Título em inglês: Comparative study between the popular and cultured norms of the Portuguese of Vitória da Conquista: nominal number agreement

Palavras-chave em inglês: Nominal number agreement. Comparative study. Standard and Substandard Portuguese. Theory of Variation and Change. Vitória da Conquista.

Área de concentração: Linguística

Titulação: Mestre em Linguística

Banca examinadora: Prof. Dr. Jorge Augusto Alves da Silva (Presidente-Orientador); Profa. Dra. Valéria Viana Sousa (Coorientadora-UESB); Profa. Dra. Cristiane Namiuti Temponi (UESB); Profa. Dra. Alícia Dunha Lose (UFBA)

Data da defesa: 20 de fevereiro de 2015

Programa de Pós-Graduação: Programa de Pós-Graduação em Linguística.

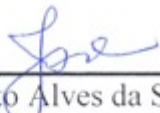
GILBERTO ALMEIDA MEIRA

**ESTUDO COMPARATIVO ENTRE AS NORMAS POPULAR E CULTA DO
PORTUGUÊS DE VITÓRIA DA CONQUISTA: CONCORDÂNCIA NOMINAL DE
NÚMERO**

Dissertação apresentada ao Programa de Pós-Graduação em Linguística (PPGLIN) da Universidade Estadual do Sudoeste da Bahia (UESB), como requisito parcial e obrigatório para obtenção do título de Mestre em Linguística.

Data da aprovação: 20 de fevereiro de 2015.

BANCA EXAMINADORA



Prof. Dr. Jorge Augusto Alves da Silva (UESB)
(Orientador)



Prof. Dra. Valéria Viana Sousa (UESB)
(Coorientadora)



Prof. Dra. Cristiane Namiuti Temponi (UESB)



Prof. Dra. Alicia Duha Lose (UFBA)

AGRADECIMENTOS

Ao fazer um exame de consciência sobre minha trajetória de vida, não tenho como olvidar os momentos bons e os momentos ruins que marcaram a minha existência. Alegrias e tristezas alternaram o meu caminhar. E ao longo desse movimento quase que dialético, não faltaram aquelas pessoas de corações fraternos e de mentes lúcidas, no pensar e no discernir, que se mostraram sempre solícitas em regar o meu ser com os mais sublimes conselhos e ensinamentos. Por isso, não posso deixar de agradecer, inicialmente, a minha família, sobretudo na pessoa de minha querida mãe **Almerinda**, mais conhecida, carinhosamente, como Teté. O seu jeito peculiar de se relacionar com os filhos, priorizando sempre o diálogo como princípio norteador de qualquer advertência ou aconselhamento, cativa-nos cada vez mais de modo que tal postura tem sido, para mim, modelo exemplar a ser seguido no meu núcleo familiar doméstico. Além do mais, ela sempre me incentivou a alçar voos mais altos na minha formação acadêmica, apoiando-me em minhas decisões profissionais.

Ao meu pai, **Gilberto**, por confiar sempre no potencial de seus filhos, nunca deixando de se orgulhar de cada um deles, sem nenhuma distinção.

A minha vovó **Estelina**, a quem chamo carinhosamente de Dinha, e que considero minha segunda mãe, pelas inúmeras e inúmeras vezes que esteve ao meu lado mostrando-me, por meio de sua simplicidade e humildade, os caminhos menos íngremes a serem trilhados.

A minha querida esposa **Mírian** por estar sempre ao meu lado, incentivando-me a seguir a diante em meus estudos, abdicando de seus sonhos e projetos para que os meus pudessem ser concretizados.

Aos meus queridos filhinhos **Daniel** e **Davi**, anjinhos que Deus colocou em minha vida para mostrar que o espetáculo de minha existência só atinge a sua plenitude temporal com a realização plena deles no que se refere à formação moral, ética e religiosa de ambos.

Aos meus queridos irmãos **Gilvan**, **Neirivan**, **Gildenor** e **Euzínio** que sempre acreditaram na minha capacidade de concretização dos meus projetos no âmbito de minha profissão.

As minhas queridas irmãs **Kátia**, **Milena**, **Mileide** e **Joana** que herdaram de nossa mãe a arte de cultivar os valores morais, éticos e religiosos mais nobres para uma convivência harmoniosa e pacífica.

Aos meus inúmeros sobrinhos e sobrinhas que encantam a todos por meio da alegria e da espontaneidade.

Aos meus cunhados e cunhadas por tornar a nossa convivência uma verdadeira celebração à amizade.

A todos os meus tios e tias pelos profundos vínculos de cordialidade, confiança e respeito que sempre nos uniram. Em especial, à tia **Dina** a quem agradeço por todo apoio concedido a mim, a minha esposa e meus filhos em relação a nossa estada em Vitória da Conquista durante os dois anos em que aí residimos.

Aos meus primos e primas que sempre nutriram por mim uma grande amizade e companheirismo.

Aos meus colegas professores de Manoel Vitorino pela saudável convivência profissional.

À Secretaria Estadual de Educação e à Secretaria Municipal de educação de Manoel Vitorino por me conceder o afastamento com ônus das minhas funções docentes, no período de dois anos, a fim de que eu pudesse me dedicar, exclusivamente, à pesquisa.

À Universidade Estadual do Sudoeste da Bahia – UESB por proporcionar o acesso ao curso de mestrado acadêmico em linguística.

À coordenação do PPGLin, Programa de Pós-Graduação em Linguística, por primar sempre pela qualidade da pesquisa.

Aos professores do PPGLin pelas inestimáveis contribuições. Em especial, à professora **Cristiane Namiuti Temponie** a minha Coorientadora, professora **Valéria Viana Sousa** pela disponibilidade em ler meu texto e sugerir, na Banca de Qualificação, alguns reajustes para o aprofundamento da pesquisa.

Aos colegas do mestrado, em especial **Vânia Raquel**, que compartilhou comigo dúvidas, incertezas e, ao mesmo tempo, confiança de que iríamos conseguir cumprir o programa do mestrado.

Ao Grupo de Pesquisa em Linguística Histórica e em Sociofuncionalismo da UESB, bem como aos graduandos do curso de Letras, membros da iniciação científica, por contribuírem, significativamente, com a audição e a transcrição das fitas gravadas durante as entrevistas.

Ao meu orientador, professor **Dr. Jorge Augusto Alves da Silva**, pesquisador de raro conhecimento e sabedoria peculiar que encanta e contagia a todos. A ele, um agradecimento especial pela disponibilidade e solicitude em dirimir dúvidas, sugerir maior objetividade nas discussões e, acima de tudo, por não hesitar em momento algum em assumir uma atitude colaborativa com a nossa pesquisa, apesar de sua intensa atividade intelectual.

E, finalmente, a **Deus**, ser sagrado e inefável, por não permitir que a chama da fé se apague jamais de minha existência.

RESUMO

Nesta pesquisa, investigamos o fenômeno da variação na concordância nominal de número no português culto de Vitória da Conquista por meio de um registro sistematizado da fala dos informantes à luz da Teoria da Variação e Mudança. Objetivamos, com isso, realizar um estudo comparativo entre os dados por nós analisados e os dados analisados por Guimarães (2014), referentes ao português popular. Para uma melhor compreensão do funcionamento do objeto pesquisado, julgamos relevante controlar as seguintes variáveis estruturais: posição linear do constituinte, posição do constituinte com referência ao núcleo do sintagma nominal (doravante SN), classe gramatical do constituinte e saliência fônica; assim como as variáveis sociais: faixa etária, sexo e escolaridade. Os dados da análise foram extraídos dos *corpora* PPVC e PCVC (Corpus do Português Popular de Vitória da Conquista e Corpus do Português Culto de Vitória da Conquista, respectivamente), constituídos pelos Grupos de Pesquisa em Linguística Histórica e em Sociofuncionalismo da Universidade Estadual do Sudoeste da Bahia – UESB. Utilizamos amostras de fala de 12 (doze) informantes do português culto cujos resultados foram comparados com os dados de 12 (doze) falantes do português popular. A análise quantitativa dos nossos dados envolveu um total de 2.205 constituintes sintagmáticos nominais submetidos ao programa estatístico Goldvarb. Já em relação ao português popular, Guimarães (2014) analisou 2.979 dados. Em face da atuação das variáveis linguísticas e extralinguísticas sobre o fenômeno por nós investigado, julgamos necessário proceder à análise a partir daquelas mesmas variáveis que foram selecionadas pelo programa como significativas, do ponto de vista estatístico a fim de chegarmos a resultados muito mais conclusivos. Assim sendo, nossa análise recaiu sobre as seguintes variáveis independentes: posição linear do constituinte, posição do constituinte com referência ao núcleo do SN, classe gramatical do constituinte, saliência fônica, sexo, escolaridade e faixa etária. Esperamos, com a presente pesquisa, mostrar como é realizada a concordância nominal de número no SN entre informantes da norma culta na comunidade de fala de Vitória da Conquista, bem como evidenciar, com base na análise das variáveis linguísticas e extralinguísticas, quais delas mais favorecem a aplicação da regra formal da concordância nominal de número.

PALAVRAS-CHAVE

Concordância nominal de número. Estudo comparativo. Normas popular e culta. Teoria da Variação e Mudança. Vitória da Conquista.

ABSTRACT

This research investigates the phenomenon of the variation in the nominal number agreement in Standard and Substandard Portuguese in Vitória da Conquista – BA, Brazil, using the Theory of Variation and Change. We aim, therefore, perform a comparative study of the data we reviewed and analyzed by Guimarães (2014) for the Substandard Portuguese. For a better understanding of how the researched object, we control the following structural variables: linear position of constituent, constituent of the position with reference to the core of SN, grammatical class of constituent and phonic saliency ; as well as social variables: age, sex and education. The data analysis were extracted from the *corpora* PCVC and PCVC (*Corpus* Português Culto de Vitória da Conquista and Português Popular de Vitória da Conquista) consisting of the Research Groups in Historical Linguistics and Sociofuncionalismo the State University of Southwest Bahia - UESB. We use speech samples of twelve (12) Standard Portuguese informants whose results were compared with data from twelve (12) speakers of Substandard Portuguese. Quantitative analysis of our data involved a total of 2,205 nominal phrasal constituents subject to statistical program Goldvarb. Regarding the popular Portuguese, Guimarães (2014) analyzed 2979 data. Given the role of linguistic and extralinguistic variables on the phenomenon investigated by us, we examine from those same variables that have been selected by the program as significant from a statistical point of view in order to reach more conclusive results. Thus, our analysis found on the following variables: linear position of constituent, constituent of the position with reference to the core of SN, part of speech of the constituent, phonics saliency, gender, education and age. We hope, with this research, show as performed the nominal number agreement between informants in SN of cultural norms in the community of speech Vitoria da Conquista, and evidence, based on the analysis of linguistic and extralinguistic variables, which of them more favor the application of formal rule nominal number agreement.

KEYWORDS

Nominal number agreement. Comparative study. Standard and Substandard Portuguese.

Theory of Variation and Change. Vitoria da Conquista.

LISTA DE GRÁFICOS E QUADROS

Quadro 1 – Características do português culto.....	37
Quadro 2 – Línguas faladas pelos indígenas que habitaram a região de Vitória da Conquista.....	91
Quadro 3 – Informantes do português culto em Vitória da Conquista.....	94
Quadro 4 – Grupo de fatores da variável posição linear do constituinte.....	98
Quadro 5 – Grupo de fatores da variável posição do constituinte com referência ao núcleo.....	99
Quadro 6 – Grupo de fatores da variável classe gramatical do constituinte.....	100
Quadro 7 – Grupo de fatores da variável saliência fônica.....	101
Gráfico 1 – Porcentagem de concordância no sintagma nominal no Português culto de Vitória da Conquista.....	106
Gráfico 2 - Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável posição linear do constituinte no PCVC.....	109
Gráfico 3 – Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável posição linear do constituinte no PPVC e no PCVC.....	112
Gráfico 4 - Variável saliência fônica no português culto de Vitória da Conquista.....	116
Gráfico 5 – Variável saliência fônica no PPVC e no PCVC.....	118
Gráfico 6 – Marcas de plural nos constituintes do SN, segundo a variável classe gramatical.....	120
Gráfico 7 – Variável faixa etária no PCVC.....	121
Gráfico 8 – Variável faixa etária no PPVC e no PCVC.....	123
Gráfico 9 – A aplicação da regra de concordância nominal de número na variável sexo no PCVC.....	124
Gráfico 10 – A aplicação da regra de concordância nominal de número na variável sexo no PPVC e no PCVC.....	125
Gráfico 11 – Variável escolaridade no PCVC.....	127

LISTA DE TABELAS

Tabela 1 – População do Brasil por etnia do século XVI ao XIX.....	21
Tabela 2 – População brasileira por etnias branca e não-branca.....	26
Tabela 3 – Distribuição da população rural e urbana do município de Vitória da Conquista (1940 a 1970).....	87
Tabela 4 – Variável dependente: concordância nominal de número no português culto ...	106
Tabela 5 – Variável dependente: concordância nominal de número no PPVC e no PCVC..	107
Tabela 6 - Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável posição linear do constituinte no PCVC.....	109
Tabela 7 – Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável posição linear do constituinte no PPVC e no PCVC.....	112
Tabela 8 – Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável posição do constituinte com referência ao núcleo no PCVC.....	113
Tabela 9 – Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável saliência fônica no PCVC.....	115
Tabela 10 – Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável saliência fônica no PPVC e no PCVC.....	116
Tabela 11 – Frequência e percentual da variante “marcas de plural nos constituintes do SN”, segundo a variável classe gramatical.....	119
Tabela 12 – Aplicação da regra de concordância nominal de número na variável faixa etária no PCVC.....	121
Tabela 13 – Aplicação da regra de concordância nominal de número na variável faixa etária no PPVC e no PCVC.....	122
Tabela 14 – Aplicação da regra de concordância nominal de número na variável sexo no PCVC.....	123
Tabela 15 – Aplicação da regra de concordância nominal de número na variável sexo no PPVC e no PCVC.....	124
Tabela 16 – Variável escolaridade no PCVC.....	127
Tabela 17 – A aplicação da regra de concordância nominal de número na variável escolaridade no PPVC.....	128

SUMÁRIO

INTRODUÇÃO.....	13
1 A FORMAÇÃO DO PORTUGUÊS BRASILEIRO.....	17
1.1 O PORTUGUÊS POPULAR DO BRASIL	17
1.1.1 O Contato entre os Povos	19
1.1.1.1 A participação dos indígenas na formação sócio-histórica do Brasil	20
1.1.1.2 A participação dos africanos na formação sócio-histórica do Brasil	23
1.1.2 Teorias Explicativas Sobre a Formação do Português Popular do Brasil.....	26
1.1.2.1 Deriva linguística	27
1.1.2.2 Transmissão linguística irregular	30
1.2 O PORTUGUÊS CULTO DO BRASIL.....	33
1.3 A POLARIZAÇÃO SOCIOLINGÜÍSTICA BRASILEIRA.....	39
2 FUNDAMENTAÇÃO TEÓRICA	44
2.1 ASOCIOLINGÜÍSTICA VARIACIONISTA.....	44
2.1.1 O advento da sociolinguística.....	44
2.1.2 Teoria da variação e mudança.....	46
2.1.3 Princípios teóricos fundamentais da sociolinguística	48
2.2 ESTUDOS SOBRE A VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO PORTUGUÊS DO BRASIL	52
3 METODOLOGIA.....	81
3.1 PROCEDIMENTOS METODOLÓGICOS DA PESQUISA.....	81
3.1.1 Perfil sócio-histórico da comunidade de fala de Vitória da Conquista.....	82
3.1.1.1 O acesso ao letramento na comunidade de fala de Vitória da Conquista	85
3.1.2 A pesquisa de campo.....	87
3.1.3 Constituição do <i>corpus</i>.....	88
3.1.4 Perfil dos informantes.....	89
3.1.5 As entrevistas.....	90
3.1.6 Processo de coleta e armazenamento de dados	91
3.1.6.1 Critérios da transcrição	91
3.1.7 Levantamento de dados.....	92
3.2 A VARIÁVEL DEPENDENTE E AS VARIÁVEIS INDEPENDENTES.....	92

3.2.1 Variáveis linguísticas ou estruturais	93
3.2.1.1 Posição linear do constituinte	93
3.2.1.2 Natureza do constituinte e sua posição em relação ao núcleo do SN	94
3.2.1.3 Classe gramatical do constituinte.....	95
3.2.1.4 Saliência fônica	96
3.2.1 Variáveis extralinguísticas ou sociais	97
3.2.1.1 Sexo.....	97
3.2.1.2 Escolaridade	98
3.2.1.3 Faixa etária.....	98
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS.....	99
4.1 A VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO SINTAGMA NOMINAL NO PORTUGUÊS POPULAR E CULTO DE VITÓRIA DA CONQUISTA..	100
4.1.1 Posição linear do constituinte	103
4.1.2 Posição do constituinte com referência ao núcleo.....	107
4.1.3 Saliência fônica.....	109
4.1.4 Classe gramatical do constituinte.....	112
4.1.5 Variável social: <i>faixa etária</i>	114
4.1.6 Variável social: <i>sexo</i>	116
4.1.7 Variável social: <i>escolaridade</i>	119
CONCLUSÃO.....	122
REFERÊNCIAS.....	124
APÊNDICE.....	129
ANEXO	130

INTRODUÇÃO

Em 2011, presenciamos, estupefatos, uma verdadeira onda de revolta e indignação por parte de muitas pessoas desinformadas, linguisticamente, com relação ao famigerado episódio do livro de português do MEC,¹ noticiado pela mídia como “o livro que ensinava a falar errado”. O fato do MEC distribuir um livro que reconhecia como legítimas construções típicas da fala popular, como *Nós pega os peixe*, por exemplo, foi o suficiente para despertar em muitos setores da sociedade atitudes explícitas de preconceito linguístico, denunciando com isso o pesado estigma social que se abate sobre a fala popular. Parece que a divisão linguística do Brasil em uma norma culta e uma norma popular, reconhecidas como duas modalidades legítimas de expressão da língua, ainda não foi bem interpretada no contexto atual. A sociedade, de um modo geral, ainda não se deu conta de que

o reconhecimento da diversidade linguística é parte de uma orientação inclusiva e de respeito a diferença que é a base da concepção atual de cidadania. E não entra em contradição com o ensino da norma linguística valorizada socialmente. Não é difícil compreender que o ensino dessa norma padrão não deve buscar a disseminação e menosprezo das formas linguísticas características das comunidades humildes de boa parte de nossos alunos. Essa disseminação constitui como um ato de violência simbólica desnecessária e inaceitável².

Além do mais, a mudança na concepção de ensino de língua portuguesa no Brasil já ganhara uma nova diretriz a partir da publicação dos Parâmetros Curriculares Nacionais, em 1997. No texto relacionado à disciplina de língua portuguesa, dedicado às séries iniciais do ensino fundamental, consta de forma bastante explícita a seguinte orientação:

A língua portuguesa, no Brasil, possui muitas variedades dialetais. Identificam-se geográfica e socialmente as pessoas pela forma como falam. Mas há muitos preconceitos decorrentes do valor social relativo que é atribuído a diferentes modos de falar: é muito comum considerar as variedades linguísticas de menor prestígio como inferiores ou erradas. O problema do preconceito disseminado na sociedade em relação às falas dialetais deve ser enfrentado, na escola, como parte do objetivo educacional mais amplo de educação para o respeito à diferença. Para isso, e também para poder ensinar Língua Portuguesa, a escola precisa livrar-se de alguns mitos: o de que existe uma única forma “certa” de falar – a que se parece com a

¹ Os críticos de plantão não se deram o trabalho, em momento algum, de ler integralmente a obra didática, pois se assim o fizessem se dariam conta, certamente, de que o livro continha apenas um capítulo introdutório sobre a variação linguística no qual se referia a divisão linguística do país entre uma norma culta e uma norma popular, reconhecidas como duas modalidades legítimas de expressão da língua, embora a norma popular fosse objeto do preconceito linguístico. No mais, toda a obra primava pelo ensino da norma padrão como condição essencial para a inserção do indivíduo na sociedade letrada.

² Anotação de trecho da palestra proferida pelo professor Dr. Dante Lucchesi na 25ª Jornada Nacional do GELNE – Grupo de Estudos Linguísticos e Literários do Nordeste, na Universidade Federal do Rio Grande do Norte, realizada de 01 a 03 de outubro de 2014.

escrita – e o de que a escrita é o espelho da fala – e, sendo assim, seria preciso “consertar” a fala do aluno para evitar que ele escreva errado. Essas duas crenças produziram uma prática de mutilação cultural que, além de desvalorizar a forma de falar do aluno, tratando a sua comunidade como se fosse formada por incapazes, denota desconhecimento de que a escrita de uma língua não corresponde inteiramente a nenhum de seus dialetos, por mais prestígio que um deles tenha em um dado momento histórico. A questão não é falar certo ou errado, mas saber qual forma de fala utilizar, considerando as características do contexto de comunicação, ou seja, saber adequar o registro às diferentes situações comunicativas. É saber coordenar satisfatoriamente o que falar e como fazê-lo, considerando a quem e por que se diz determinada coisa. É saber, portanto, quais variedades e registros da língua oral são pertinentes em função da intenção comunicativa, do contexto e dos interlocutores a quem o texto se dirige. A questão não é de correção da forma, mas de sua adequação às circunstâncias de uso, ou seja, de utilização eficaz da linguagem: falar bem é falar adequadamente, é produzir o efeito pretendido (MEC, 2001, p. 26).

Esse parecer advindo diretamente de um documento oficial produzido pelo Ministério da Educação em momento algum foi considerado nas discussões unilaterais promovidas pela mídia, na ocasião. Ao contrário, tentaram esvaziar qualquer reflexão de natureza científica acerca das pesquisas linguísticas em consonância com o posicionamento do livro em relação ao tema da variação linguística.

Ora, a frase *Nós pega os peixe*, que foi uma das que gerou a virulenta ofensiva contra o livro “Por uma Vida Melhor”, está estruturada em torno de dois sintagmas: verbal (*Nós pega*) e nominal (*os peixe*). O fato é que ambos os sintagmas encontram-se no plano morfossintático da língua portuguesa e é justamente aí nesse fenômeno que se torna mais saliente a distinção entre a concordância formal e a não concordância. Logo, qualquer mentalidade plasmada pelo preconceito linguístico acaba reagindo negativamente ao testemunhar de forma ocular ou auricular ocorrências sintagmáticas como a supracitada.

Em nosso trabalho, focalizamos a pesquisa na variação na concordância nominal de número no SN no português popular e culto de Vitória da Conquista, buscando realizar um estudo comparativo entre essas duas modalidades do português brasileiro. Numa atitude eminentemente investigativa, consideramos imprescindível, para entender o fenômeno em questão, romper com a perspectiva normativa segundo a qual a marcação de plural no SN se constitui em uma regra categórica já que a gramática tradicional prescreve para a norma padrão a devida concordância em número dos artigos, pronomes, numerais e adjetivos com os substantivos a que se referem. Na verdade, filiamo-nos aos diversos estudos sobre a concordância nominal de número já realizados no Brasil que comprovam tratar-se de um fenômeno variável, porém não aleatório, condicionado por fatores estruturais e sociais. São essas variáveis que melhor explicam, por exemplo, uma ocorrência como *os peixe*.

Para a pesquisa empírica envolvendo a variável dependente *concordância nominal de número*, foram utilizadas um total de 24 (vinte e quatro) entrevistas, sendo 12 (doze) de falantes do português culto, por nós analisadas, e 12 (doze) de falantes do português popular, analisadas por Guimarães (2014), estratificados de acordo com o sexo: masculino e feminino; duas faixas etárias: Faixa I (20 a 35 anos) e Faixa II (36 a 50 anos) e o nível de escolaridade: informantes que possuíam, no máximo, 5 anos de escolarização (falantes do PP) e informantes com mais de 11 (onze) anos de formação escolar (falantes do PC), considerando, nesse último caso, a formação profissional técnica e o nível superior. Após a rodada estatística com as variáveis independentes: *posição linear do constituinte*, *posição do constituinte com referência ao núcleo do SN*, *classe gramatical do constituinte*, *saliência fônica*, além da *faixa etária*, *sexo* e *escolaridade*; ressaltamos que o programa Goldvarb selecionou como relevantes do ponto de vista estatístico, pela ordem de seleção: *posição linear do constituinte*, *escolaridade*, *faixa etária*, *posição do constituinte com referência ao núcleo*, *sexo* e *saliência fônica*. A variável *classe gramatical* não foi selecionada pelo programa.

Pensando numa abordagem mais didática sobre a pesquisa por nós realizada, julgamos relevante dividir esse trabalho em quatro capítulos. Assim, logo no primeiro, discutimos a formação do português brasileiro, considerando os processos históricos envolvendo a origem do português popular, bem como o antecedente histórico do português culto. Discorremos, igualmente, sobre a polarização sociolinguística brasileira cujos fundamentos remetem-nos a uma reflexão mais profunda acerca da organização sócio-histórica do país.

No segundo capítulo, fazemos uma breve consideração da teoria que embasa nosso estudo, a sociolinguística variacionista de cunho laboviano. Apresentamos também uma revisão bibliográfica acerca de alguns estudos sobre a variação na concordância nominal de número no português do Brasil.

Apresentamos, no terceiro capítulo, os procedimentos metodológicos da pesquisa, focalizando aspectos relativos ao perfil sócio-histórico da comunidade de fala de Vitória da Conquista, à pesquisa de campo, à constituição do *corpus*, ao perfil dos informantes, às entrevistas, ao processo de coleta e armazenamento de dados, ao levantamento de dados, bem como à variável dependente e às variáveis independentes.

Tratamos da análise e discussão dos dados no quarto capítulo, apresentando as variáveis independentes selecionadas significativas pelo programa Goldvarb.

E, por fim, na conclusão, julgamos relevante apresentar uma síntese dos resultados por nós obtidos, dando ênfase, sobretudo, aqueles pontos mais diretamente relacionados a uma abordagem contrastiva entre as normas popular e culta do português de Vitória da Conquista.

Pontos esses oriundos, justamente, do índice de variação na indicação de pluralidade entre os elementos flexionáveis do SN.

Esperamos, com esse estudo, contribuir para o conhecimento da realidade sociolinguística de Vitória da Conquista, tendo em vista a afirmação do município como mais uma área dialetal brasileira com um fecundo campo sociolinguístico a ser cada vez mais explorado.

1 A FORMAÇÃO DO PORTUGUÊS BRASILEIRO

Quando as primeiras embarcações portuguesas aportaram na costa brasileira, em 1500, a euforia dos tripulantes e comandantes com o feito, certamente, se traduziu na emissão das primeiras palavras em língua portuguesa em solo brasileiro. Era a matriz linguística europeia se impondo, timidamente, em uma enorme nação cujos legítimos habitantes, espalhados aos milhares, em diferentes regiões, falavam diversas línguas aborígenes. A partir de 1549, com a chegada dos primeiros escravos trazidos do continente africano, inaugura-se mais uma etapa na configuração linguística do país. É do encontro dessa diversidade de línguas e de culturas que emerge o português brasileiro, com peculiaridades sócio-históricas que além de distingui-lo do português de Portugal, ensejam o desabrochar de duas normas³: popular e culta. A primeira, extremamente arraigada no contexto cultural, histórico e linguístico dos primeiros séculos do Brasil, era falada pela grande maioria das pessoas excluídas da cultura letrada. O que não é muito diferente da realidade atual. A segunda, restrita a uma minoria que era herdeira dos valores europeus.

Nesse sentido, acreditamos que recontar os percursos da constituição histórica das duas principais normas do Português do Brasil nos ajudará a compreender melhor o fenômeno da variação da concordância de número no Sintagma Nominal, bem como compreender as tendências apontadas pela quantificação dos dados preliminares alcançados por nosso recorte. Assim, podemos traçar não apenas um quadro de variação, mas compreender o curso da mudança linguística no português popular e no português culto.

Para uma melhor compreensão do desenrolar desses fatos morfossintáticos envolvendo a formação do português brasileiro, dividimos este primeiro capítulo em três seções: o português popular do Brasil, o português culto e a polarização sociolinguística brasileira. As duas primeiras seções constam de subseções nas quais serão discutidas, de forma mais pontual, cada aspecto relacionado ao tema, o que permitirá uma abordagem muito mais didática.

1.1 O PORTUGUÊS POPULAR DO BRASIL

As discussões em torno da formação do português popular brasileiro têm sido bastante acirradas nas últimas décadas entre os pesquisadores que se debruçam sobre a

³ No presente estudo, tomamos “norma” a partir do conceito reelaborado por Eugenio Coseriu (1979) o qual trata a linguagem numa visão tripartida, a saber: sistema, norma e fala.

realidade sociolinguística do país. Motivados pela constatação de que a língua reflete as estruturas sociais dos falantes – fato ignorado pelo estruturalismo, como lembra Lucchesi (1998) – os estudiosos acabaram transitando por distintas teorias explicativas acerca da origem do português popular do Brasil (doravante, também PPB), haurindo daí os elementos necessários para fundamentar determinado posicionamento. Dentre as teorias envolvidas nos debates sobre o PPB, merecem destaque, no presente estudo, a deriva secular e a transmissão linguística irregular, cujos representantes mais notáveis são, sem dúvida, Anthony Julius Naro, Maria Marta Pereira Scherre (2007)⁴, Dante Lucchesi (2000), Jorge Augusto Alves da Silva (2005), entre outros. Ambas as teorias serão aqui discutidas de forma mais aprofundada em função da necessidade de apresentar possíveis evidências da vinculação da variação na concordância plural no sintagma nominal a uma dessas abordagens teóricas.

É mister contextualizar, preliminarmente, o processo de formação histórica e social do PPB, tendo em vista encontrar elementos que ajudem a compreender as peculiaridades dessa variedade, sobretudo, no que se refere aos fenômenos de variação. Nesse sentido, parece oportuno trazer à baila algumas considerações relativas ao termo popular. E para tanto, nada mais esclarecedor do que o trabalho de Silva (2005). Em sua tese de doutorado, intitulada “A Concordância Verbal de terceira pessoa do plural no português popular do Brasil: um panorama sociolinguístico de três comunidades do interior do estado da Bahia”, o autor discute o conceito do termo “popular” a partir de definições tanto numa perspectiva etimológica quanto histórica para então caracterizar a variedade de língua considerada popular. Ele afirma que “o termo popular sempre esteve associado à oposição ao culto, ao aristocrático, à elite”. E lembra que “ao nomearmos uma língua de popular, estamos assumindo que há uma diferença perceptível na forma de falar do povo e do não-povo ou elite ou camada culta” (SILVA, 2005, p. 10).

Como critério para definição do termo, numa investigação de natureza sociolinguística, o pesquisador utiliza a palavra popular para fazer referência a dois grupos sociais, a saber: “àquela parcela da população brasileira que não passou pelo processo de letramento e àquela parcela constituída por aqueles que frequentaram os bancos escolares por quatro anos quer na infância ou quando atingiram a maturidade” (SILVA, 2005 p. 13). É nessa perspectiva conceitual que este trabalho, também, se insere, visto que o fenômeno da variação de pluralidade nos constituintes do SN, aqui pesquisado, envolve, também, falantes com esses mesmos perfis escolares.

⁴ Em *Origens do Português Brasileiro*, os autores assumem de forma inequívoca a noção de deriva linguística com base em Edward Sapir, na obra *A linguagem: introdução ao estudo da fala*.

A relação intrínseca entre língua popular e classe social, fato consensual entre os sociolinguistas, encontra sua expressão máxima no conhecimento da formação sócio-histórica dessa variedade vernacular. Essa constatação pode ser respaldada ainda nas palavras do autor, quando o mesmo afirma que

[...] a caracterização da língua popular do Brasil está diretamente ligada à classe social de seus falantes e a classe social do falante está relacionada ao seu contexto sócio-histórico. cremos, portanto, que o falante do português popular do Brasil possua como língua vernácula, nos termos enunciados por Labov (1972), uma variedade de língua que reflete a constituição histórica das classes menos favorecidas do Brasil, perceptível no nível social, econômico e cultural dos utentes da variedade popular tanto no meio rural quanto no meio urbano (SILVA, 2005, p. 15).

Depreende-se, igualmente, das palavras de Silva (2005), que a análise do contexto da sócio-história dos falantes do PPB só atinge a sua completude quando se leva em consideração aspectos relativos a origem social desses agentes formadores da realidade sociolinguística brasileira. Portanto, o âmbito discursivo envolvendo a gênese do PPB tem um alcance muito maior do que se supunha, daí a importância em explorar o tema numa perspectiva mais ampla, trazendo para a reflexão desde questões atinentes ao contato social entre os agentes quanto o processo de formação da sociedade brasileira, de um modo geral.

1.1.1 O Contato entre os Povos

A presença do elemento indígena e africano na constituição da história social do Brasil por si só já é um indicativo mais que eloquente do caráter vocacional do país em abarcar a diversidade linguística. A existência de diversas línguas indígenas faladas pelos nativos mais a grande quantidade de línguas e dialetos que vieram com os escravos oriundos de várias regiões do continente africano fizeram do Brasil um celeiro linguístico bastante privilegiado do ponto de vista das variedades. A língua portuguesa aqui transplantada pelo colonizador encontrou as condições necessárias para seguir o curso da mudança e da variação, afastando-se da matriz europeia. E não há dúvida de que as peculiaridades observadas no português popular, seja de natureza fonética, morfológica, semântica ou sintática, se devem muito ao contato linguístico e social estabelecido por aqui, no qual os agentes formadores da sócio-história brasileira: índios, negros e europeus protagonizaram os momentos de interação verbal, que certamente foram tímidos no início, mas se intensificaram com o passar do tempo. A esse respeito, Antonino (2012) afirma que:

No período de formação da língua falada em terras brasileiras, o português transplantado da Europa, ocorreu um convívio intenso de falantes de diversas outras línguas. Era um número muito grande de negros trazidos forçosamente de diversas regiões africanas, índios, autóctones, também de diferentes tribos, falantes de diferentes línguas e, ainda, os brancos portugueses, também de classes sociais variadas. A convivência tornou-se inevitável, e muitas vezes indispensável, pois habitavam a mesma terra e havia, entre eles, relações sociais que exigiam que se comunicassem minimamente (ANTONINO, 2012, p. 22).

Nessa mesma linha de raciocínio, Lucchesi (2009) destaca o fato de que:

O contato dos colonizadores portugueses com milhões de aloglotas, falantes de mais de mil línguas indígenas autóctones e de cerca de duzentas línguas que vieram na boca de cerca de quatro milhões de africanos trazidos para o país como escravos, é, sem sombra de dúvida, o principal parâmetro histórico para a contextualização das mudanças linguísticas que afetaram o português brasileiro (LUCCHESI, 2009, p. 41).

Nas discussões a seguir, será enfatizada a participação efetiva dos indígenas e dos negros africanos enquanto agentes formadores da sociedade brasileira e de sua realidade sociolinguística. O elemento branco, europeu, não será, obviamente, ignorado nesse estudo. A análise de sua participação se dará em momento oportuno, sobretudo, quando se procurará evidenciar situações de contraste entre os colonizadores e os povos subjugados: nativos e escravos.

1.1.1.1 A participação dos indígenas na formação sócio-histórica do Brasil

A chegada das primeiras levas de portugueses colonizadores ao Brasil, em 1500, inaugurou uma nova era na história da civilização indígena local. A exuberância da fauna e da flora brasileira não foi suficiente para conter o ímpeto explorador dos colonos, que movidos por interesses mercantilistas não hesitaram em momento algum em fazer uso da força para escravizar os índios. Num primeiro momento, tendo em vista o reconhecimento da região, julgaram mais prudente utilizar o artifício conhecido como escambo para conquistar a simpatia do indígena. Com isso, foi possível manter um clima aparentemente amistoso que se refletiu na colaboração do nativo na realização de alguns serviços, como a exploração do pau-brasil, por exemplo. Mas essa realidade não perduraria por muito tempo. A natureza prodigalizou aos nativos tudo que lhes era necessário para sobreviver. Por isso mesmo, o senso de liberdade lhes era quase que congênito. Porém, ao se sentirem acossados pelo invasor e verem ameaçada a soberania de seu território, os aborígenes reagiram contra o instinto selvagem do colonizador, dando início a um segundo momento em que a convivência

entre eles se tornaram bastante instável. O final dessa história já é por demais conhecido por todos: milhões de indígenas dizimados em um intervalo de três séculos. Em sua dissertação de mestrado sobre “O papel das línguas africanas na história do português do Brasil”, escrita ainda em 1991, o autor Alberto Mussa apresenta uma tabela elaborada por ele na qual é possível visualizar esse intenso declínio que houve da população autóctone, e que vale a pena reproduzi-la abaixo:

Tabela 1: População do Brasil por etnia do século XVI ao XIX

Etnias	1538-1600	1601-1700	1701-1800	1801-1850	1851-1890
Africanos	20%	30%	20%	12%	2%
Negros brasileiros	-	20%	21%	19%	13%
Mulatos	-	10%	19%	34%	42%
Branços brasileiros	-	5%	10%	17%	24%
Europeus	30%	25%	22%	14%	17%
Índios integrados	50%	10%	8%	4%	2%

Fonte: Mussa (1991, p. 163).

É fato mais que notório os esforços dos jesuítas para impedir a exploração dos indígenas, exigindo, inclusive, das autoridades portuguesas o fim da escravidão desses povos. Os apelos parecem não ter sido suficientes, pois os interesses materiais estavam muito além dos espirituais. E a consequência de tudo isso pode ser vista na tabela 1 logo acima. A exorbitante diminuição de índios a partir do século XVII, justamente no período em que a exploração das terras brasileiras foram se intensificando cada vez mais, é um testemunho incontestável de que não faltaram expedientes para dizimar os nativos. Conforme Silva (2005, p. 20), “milhares de índios foram dizimados quer por guerras movidas contra eles pelos brancos quer por doenças adventícias contra as quais não tinham defesa”. A morte de boa parte dessas populações ameríndias, de diferentes tribos, representou também a extinção de várias línguas faladas por esses povos. Tomando por base Aryon Rodrigues, Mattos e Silva (2004, p. 76) lembra que esse autor

[...] admitindo para a atualidade cerca de 180 línguas sobreviventes, considerou possível o dobro em 1500. Em artigo posterior (1993:91), com base em cálculos de distribuição de línguas no passado da humanidade como um todo, faz crescer esse número para 1.175 línguas. Antônio Houaiss (1985:100) admite verossímil o montante de 1.500 línguas.

Em meio à multiplicidade de línguas, dá para imaginar de que forma se deu o contato inicial entre colonizadores e indígenas: através de gestos, certamente. Silva (2005) afirma que essa situação vai se modificando já que “alguns homens brancos acabaram por ficar nas tribos, aprendendo-lhes a língua” (cf. p. 19) e complementa tal informação assegurando que:

Os primeiros portugueses que vieram para o Brasil eram aventureiros em busca de riquezas ou condenados pela justiça banidos para longe da corte. Tais homens estabeleciam relações matrimoniais com as índias, favorecidos pela prática do cunhadismo. Um alvará de 1775 atesta certo incentivo aos casamentos entre portugueses e índias, gerando mestiços, os mamelucos para quem a língua materna era o tupi, mas podendo ser bilíngues, adquirindo o português como segunda língua (SILVA, 2005, p. 19).

Há quase que um consenso entre os pesquisadores de que ao longo dos primeiros séculos de colonização uma língua geral tenha se tornado o principal instrumento de comunicação para o colonizador e para os jesuítas na interação com os índios. Rodrigues (2000 *apud* SILVA, 2005, P. 19) defende a existência de duas línguas gerais no período colonial: uma língua geral paulista e uma língua geral amazônica. Nessa mesma linha de pensamento, Ilari e Basso (2009) observam que:

Apesar da variedade de línguas indígenas presentes, a criação de ‘línguas gerais’ era facilitada, no Brasil, pelo fato de que as línguas nativas da costa, pertencentes em sua maioria ao tronco tupi, apresentavam uma relativa uniformidade; foi a partir dessas línguas que se criaram as línguas gerais brasileiras. Uma delas teve grande difusão na região Sudeste e continuou sendo falada em São Paulo até o início do século XX; outra, conhecida como *nheengatu* (que literalmente significa “boa língua” – foi assim batizada para estabelecer uma oposição com as línguas que ajudava a silenciar), teve difusão no Norte e ainda sobrevive em regiões circunscritas da Amazônia (ILLARI & BASSO, 2009, p. 62).

Essa realidade começa a se modificar no começo do século XVIII, a partir do momento em que as autoridades portuguesas consideraram que, para o progresso da colônia, era necessário que se consolidasse a língua portuguesa em terras brasileiras, partindo, para tanto, do ensino formal que se encontrava sob a responsabilidade dos jesuítas. A esse respeito, Noll (2008, p. 208) chama a atenção para a iniciativa do governador-geral João de Lencastro, propondo ao rei de Portugal que “fossem fundados dois seminários para moços e moças índias de quatro a doze anos” a fim de que o português fosse a única língua falada e ensinada em tais estabelecimentos. O autor lembra, também, que outras medidas foram tomadas para que o português pudesse ser fomentado na educação dos indígenas. Como é o caso de uma carta

datada de 1727, endereçada ao governador do Norte do Brasil, na qual o rei havia decretado a proibição da língua geral nas colônias e nas aldeias indígenas.

Ainda, segundo Noll (2008), parece que esse interdito não pode ser implementado naquele momento, certamente, por conta da resistência dos jesuítas que faziam uso da língua geral para catequizar o gentio. As pretensões linguísticas da administração portuguesa só começam a se concretizar de fato com a ascensão do Marques de Pombal à condição de regente absoluto do reino, como primeiro-ministro de Portugal. Uma das primeiras medidas tomadas por ele foi a expulsão dos jesuítas tanto da metrópole quanto das colônias. O que atingiu diretamente os membros da companhia de Jesus instalados no Brasil. A partir daí, não foi difícil a imposição e oficialização da língua portuguesa por aqui, até mesmo entre os índios catequizados e domesticados pelos jesuítas.

1.1.1.2 A participação dos africanos na formação sócio-histórica do Brasil

Embora a esquadra de Cabral tenha aportado em terras brasileiras no ano de 1500, o interesse em explorar vorazmente os recursos naturais do território só ganhou uma proporção ingente a partir de 1549. Foi nesse período que o então rei de Portugal, Dom João III, autorizou “a importação de 120 peças anuais por agricultor para o trabalho na lavoura de cana-de-açúcar e nos engenhos” (SILVA, 2005, p. 21). De acordo com Lopes (2001, p.41), “de início, o número de africanos no território brasileiro era menor que o de portugueses e índios, mas, no século XVIII, com a intensificação do comércio de escravos, em algumas regiões eles passaram a ser maioria”.

Parece não haver consenso, entre os estudiosos, sobre o número aproximado de escravos trazidos à força para o Brasil. A esse respeito, Lucchesi (2000) considera que, até a extinção do tráfico de africanos, os dados que se têm são um tanto pendulares, oscilando entre números díspares, o que dificulta uma estimativa aproximada em termos quantitativos. Para ele,

As estimativas do número de negros africanos trazidos para o Brasil durante cerca de três séculos de tráfico negreiro variam enormemente. Na ausência de fontes diretas, só é possível chegar a alguma cifra através de processos quantitativos indiretos. Pandiá Calógeras (1927[1957]) partiu de dados de recenseamentos da escravidão no início do século passado (isto é, XIX) para chegar a uma média anual de 55.000 escravos importados. Desse modo, apresentou um número impressionante de treze milhões e quinhentos mil escravos importados, que a maioria dos estudiosos considera exagerado. Roberto Simonsen (1937), baseando-se na produtividade e na vida média dos escravos, chegou a um número bem menor: três milhões e trezentos mil. Já Maurício Goulart (1949), que adotou um método direto, com base nas poucas informações remanescentes do tráfico, chegou a um número semelhante, que

oscilaria entre três milhões e quinhentos e três milhões e seiscientos mil. Essa é a cifra que normalmente é tomada por base nos estudos sobre a escravidão no Brasil (LUCCHESI, 2000, p. 59).

Outra questão um tanto lacunar tem a ver com a procedência dos escravos em relação ao território africano de origem, assim como, a quantidade de línguas que vieram com esses povos. A historiadora Kátia Mattoso tem sido um dos nomes mais consultados quando se trata da escravidão no Brasil. Em sua obra, *Ser escravo no Brasil* (1990, p. 22-23), a autora se baseia em três ciclos que permitem uma identificação mais segura no que tange à procedência dos africanos: o Ciclo da Guiné, no século XVI, que compreendem os uolofs, os mandigas, os sonrais, os mossis, os hauças e peuls; o Ciclo do Congo e de Angola, no século XVII e o Ciclo da Mina, do Golfo de Benin e da Costa da Mina, no século XVIII. As levas de escravos trazidos no século XIX eram, na sua grande maioria, oriundas de Angola e Moçambique.

Quanto ao número de línguas faladas por essa população cativa, parece não haver como precisar com exatidão, mas não faltam esforços de linguistas e historiadores nessa direção. Lucchesi (2009), por exemplo, considera “cerca de duzentas línguas que vieram na boca de cerca de quatro milhões de africanos” (LUCCHESI, 2009, p. 41). O historiador Luís Henrique Dias Tavares (2001) não fala em quantidade, mas destaca o fato de que era possível identificar os escravos trazidos para a Bahia por meio dos falares iorubas, ewes, jejes, fulas, tapas, ardas, calabares e malês. Ainda, de acordo com o estudioso, esses últimos além de falar, escreviam em árabe, sugerindo, com isso, o acesso à leitura do Corão.

Ao se deparar com a diversidade de línguas que vieram com os escravos de diferentes regiões do continente africano, o colonizador, temendo possíveis rebeliões e fugas, julgou mais prudente manter os falantes de línguas distintas convivendo juntos em um mesmo espaço, dificultando, dessa forma, a comunicação entre eles. Rodrigues (1983), a esse respeito, avalia a situação nessa mesma perspectiva, assegurando que o objetivo do colonizador em promover a diversidade intencional de línguas era evitar a coesão social entre grupos de escravos, mantendo-os submissos mais facilmente. Esse critério, nem um pouco humanístico, para justificar a separação de pessoas que compartilhavam uma mesma língua materna e que tiveram sua existência transplantada em um país estrangeiro reflete a lógica desumana com a qual o homem branco se deixava levar para impor o seu senhorio. O episódio parece parodiar a história da torre de babel, na qual é narrado o desejo do homem que, desobedecendo a Deus, pretendia subir aos céus escalando uma torre construída para tal fim. Como forma de castigo para tamanha audácia, Deus lança a confusão das línguas entre o povo, fazendo com que eles não conseguissem entender um ao outro, impossibilitando,

portanto, o projeto. Os colonizadores europeus se sentiam o próprio Deus. E foi movido por esse sentimento que eles fizeram da escravidão retórica de desenvolvimento econômico e progresso da nação brasileira, quando, na verdade, o resultado foi bem outro, como se pode ver nos contrastes sociais pelo país afora.

O tráfico interno tem sido outro aspecto da escravidão, no Brasil, muito explorado pelos estudiosos, mormente entre os linguistas, já que envolve a mobilidade de negros africanos no território brasileiro, fato que contribuiu, significativamente, para a difusão do português aprendido por eles nas fazendas de cana-de-açúcar. Para reforçar essa ilação, vale, por sinal, retomar a discussão anterior, que tem a ver com a convivência entre escravos falantes de línguas diferentes. Ora, considerando que o contato deles com o português era, provavelmente, muito pouco, já que “nas plantações não havia muitos portugueses” (LOPES, 2001, p. 41), imagina-se que tanto a ausência de uma língua comum quanto a precariedade, em termos interacionais, tiveram implicações na aquisição do português por parte dos africanos e de seus descendentes. Com as descobertas das minas de ouro e diamante em Minas Gerais, Mato Grosso e Goiás e o declínio da lavoura canavieira em Pernambuco, Rio de Janeiro e Bahia, verificou-se um repentino fenômeno migratório no país, em que milhares de escravos foram deslocados para as regiões, economicamente, mais prósperas. Ademais, as lavouras de café em São Paulo, Rio de Janeiro e Minas Gerais absorveram, igualmente, a mão-de-obra escrava. O resultado de tudo isso, numa perspectiva linguística, pode ser encontrado na explicação de Mattos e Silva (2000) para quem essa crescente presença de escravos em várias localidades dessa enorme nação brasileira é um indício de que esses povos, por onde passavam, propagavam o português popular aprendido nas lavouras e senzalas. A tabela II a seguir, extraída de Lobo (1996, p. 16) e que fora utilizada por Silva (2005, p. 24) e Antonino (2012, p. 24) para realçar o quadro multiétnico da sociedade brasileira, nos períodos colonial e imperial, é reproduzida aqui também a fim de estabelecer uma relação entre as etnias não-brancas e branca e o uso das variedades do português:

Tabela 2 : População brasileira por etnias branca e não-brancas

	Etnias não-brancas	Etnia branca
1558-1600	70%	30%
1601-1700	70%	30%
1701-1800	68%	32%
1801-1850	69%	31%
1851-1890	59%	41%

Fonte: LOBO, 1996, p. 16.

Os dados da tabela 2 sugerem, do ponto de vista linguístico e social, que a maior parte da população brasileira era falante do português popular. Lobo (1996) identifica as etnias não-brancas aos africanos, afrodescendentes e índios aculturados. Essa parcela da sociedade estava excluída do acesso à cultura letrada e, portanto, a margem de um contato mais efetivo com o português culto, falado pela minoria da população que compunha a etnia branca: portugueses europeus e descendentes mais europeizados.

Diante de tudo o que foi exposto até aqui, é possível pensar no quanto a formação de uma língua está relacionada à constituição de uma sociedade. Até porque, como lembra Calvet (2002, p. 12) “as línguas não existem sem as pessoas que as falam, e a história de uma língua é a história de seus falantes”. No Brasil, esses falantes compuseram, desde a época colonial, dois quadros sociais bastante distintos cuja descrição apresentada a seguir, de autoria de Silva (2005), sintetiza com perfeição a realidade social, tanto de outrora quanto do presente:

No caso brasileiro, podemos ver claramente a formação de, pelo menos, duas sociedades. Uma formada por brancos de origem lusitana e seus herdeiros, especialmente, por aqueles que se esmeraram em não só possuir os meios de riqueza, mas também de se apossarem de outras formas de poder, pela legitimação de profissões que exigissem especializações. Do outro lado, formar-se-ia uma sociedade herdeira dos negros, dos índios e dos brancos empobrecidos, cujo papel de atuação estava no exercício de atividades pouco remuneradas e com exigência de pouca especialização. Assim, a história social do Brasil reflete a história econômica das atividades praticadas no Brasil e ambas estão relacionadas à variedade de língua que um e outro lado usam como forma de expressão (SILVA, 2005, p. 24).

As discussões ulteriores vão centrar na modalidade popular da língua portuguesa do Brasil, dando ênfase, sobretudo, as teorias explicativas sobre a formação do PPB. Será debatido, do mesmo modo, o português culto, procurando traçar o perfil dos falantes relacionados, principalmente, ao letramento.

1.1.2 Teorias Explicativas Sobre a Formação do Português Popular do Brasil

Nas subseções seguintes, 1.1.2.1 e 1.1.2.2, discutiremos as teorias explicativas sobre a formação do português popular do Brasil: deriva linguística e transmissão linguística irregular, atentando-nos para o fato de que embora essas abordagens partam de perspectivas conceituais distintas, são unânimes em reconhecer que a língua reflete a hierarquia social do falante.

1.1.2.1 Deriva linguística

Nas reflexões apresentadas por Naro e Scherre sobre as origens do português popular do Brasil, o que se observa é uma forte tendência desses autores em refutar veementemente às teses de suposta pidginização e crioulização do português brasileiro, apoiando-se em evidências históricas e sociais, como sustentam ambos em sua obra: *Origens do Português Brasileiro*. Ali, eles são bastante enfáticos em defender a noção de deriva linguística de Edward Sapir segundo a qual "a língua se move ao longo do tempo num curso que lhe é próprio. Tem uma deriva" (SAPIR, 1971, p. 150). Mas adiante, esse autor vai ilustrar essa inclinação para a mudança partindo da seguinte constatação "as mudanças dos próximos séculos estão em certo sentido prefiguradas em certas tendências não perceptíveis do presente" (SAPIR, 1971, p. 155). Com isso, aqueles autores, embora tenham percebido a existência de uma enorme incompatibilidade entre o processo de formação do português popular do Brasil e a influência de um suposto crioulo de base lexical portuguesa, não atribuíram importância alguma a tal fato. Na verdade, para eles, houve uma confluência de motivos que desencadearam para a constituição do português popular, a saber: "a atração de forças de diversas origens - algumas oriundas da Europa, outras da América, outras, ainda, da África - que, juntas, se reforçaram para produzir o português popular do Brasil" (NARO & SCHERRE, 2007, p. 25). Daí, afirmarem categoricamente que

além de outras forças em interação com a deriva secular trazida da Europa, parece-nos oportuno lembrar também o papel dos índios, o das demais etnias presentes no Brasil e a contribuição pidginizante dos primeiros colonos portugueses (NARO & SCHERRE, 2007, p. 26).

Em relação a essa contribuição pidginizante, Naro e Scherre (2007) destacam o fato de que, muito antes de chegarem ao Brasil, os portugueses mantiveram contato com muitos povos que não dominavam a língua portuguesa. E tal contato se dava por meio de um sistema verbal conhecido como *sabir*. Ainda, segundo os autores, trata-se de um "sistema extremamente flexível, podendo comportar itens lexicais de diversas línguas românicas (ou até do árabe)" (NARO & SCHERRE, 2007, p. 27). Portanto, tinha características bem próximas de pidgin. Não é à toa que Naro, em um artigo intitulado *A Study on the Origins of Pidginization*, afirma que as variantes do *sabir* já possuíam "traços típicos de pidgin/crioulo", que se apresentavam de forma variável, como é o caso das formas *mim saber* em vez de *eu sei*, por exemplo.

Esses fatos levam os autores a inferirem a possibilidade dos colonizadores portugueses já terem tido uma boa experiência em termos de comunicação com estrangeiros ao chegarem em terras brasileiras.

Eles também consideram que, no período colonial brasileiro, houve um

predomínio quase total da língua geral, um *pidgin* ou coine simplificado de origem tupi [...] todos os grupos étnicos - índios de diversas tribos, europeus e africanos - sabiam se comunicar usando esse sistema tupi (NARO & SCHERRE, 2007, p. 28).

Os autores são, igualmente, enfáticos ao descartar qualquer possibilidade de ter havido no período colonial um suposto crioulo de base lexical portuguesa que seria, porventura, falado pelos escravos. Eles chegam a afirmar que

[...] é escassa qualquer evidência documentária específica quanto ao português e outras línguas faladas pelos africanos no Brasil, além de algumas menções breves de africanos que não dominavam o português, mas apenas a língua geral tupi ou línguas africanas (NARO; SCHERRE, 2007, p. 28).

Sendo assim, é perfeitamente compreensível as considerações que eles fazem acerca da configuração linguística no contexto colonial brasileiro, descrevendo-a como “uma comunidade em que as línguas dos diversos grupos se influenciavam, principalmente através do aprendizado de segundas línguas por falantes não-nativos adultos” (cf. p. 29). Mas é oportuno destacar também um texto publicado em 1620 no qual, segundo os autores, há a imitação da fala dos índios feita por um missionário. O português aí usado se assemelharia ao português de pretoda Europa e da África. De todo modo, os autores não deixam dúvida de que “uma influência direta dos africanos sobre os índios teria sido impossível no início do século XVII, devido ao predomínio da língua geral e ao menor número de africanos nessa época” (NARO; SCHERRE, 2007, p. 30).

Em *Os africanos no Brasil*, a autora Nina Rodrigues discute a possibilidade de ter havido no país dois sistemas verbais de base africana cuja formação se devia ao contato entre diversas línguas africanas, tanto entre si quanto entre elas e o português. Esses sistemas seriam um *pidgin* de base ioruba, com concentração mais ao norte do Brasil; e outro *pidgin* de base quimbundo, falado mais ao sul do país. A hipótese da autora é reforçada com a descoberta de uma “língua africana” na comunidade do Cafundó. Com os estudos ali realizados por Vogt e Fry, concluiu-se que se tratava, na verdade, de um *pidgin* de base quimbundo. Portanto, ao contrário dos *pidgins* e crioulos conhecidos por possuírem gramática africana e léxico português, em Cafundó ocorreu o oposto em relação ao *pidgin* aí descoberto.

Todos esses fatos reforçaram mais ainda a conclusão de Naro e Scherre (2007), segundo a qual

[...] parece então improvável que tenha existido no Brasil uma língua pidgin ou crioula de base lexical portuguesa associada predominantemente com a etnia afro-brasileira ou ameríndia. Tal língua era dispensável, dada a existência de outras 'línguas gerais', de bases não-européias, que já preenchiam as necessidades comunicativas da população (NARO; SCHERRE, 2007, p. 31-32).

Levando em consideração que o português popular do Brasil se caracteriza pela variação nos sistemas de flexão nominal e verbal, a variabilidade em tais sistemas tem sido, conforme Naro e Scherre (2007, p. 32), um dos fenômenos apontados como argumento para se atribuir a influência de línguas africanas no processo de formação do português brasileiro. Os autores, inclusive, sinalizam para algumas particularidades envolvendo as línguas crioulas e africanas que supostamente poderiam se constituir em prova cabal a favor da origem crioula da variação nos mecanismos de concordância. A esse respeito, esses autores, ao contrário, postulam que, por força de uma deriva secular, a língua portuguesa já possuía “o embrião do novo sistema mais analítico, antes mesmo de sair da Europa” (NARO; SCHERRE, 2007, p. 32). Para reforçar mais ainda essa perspectiva, os autores recorrem ao estudo de Leite de Vasconcellos, mormente, no que se refere à queda do -s no português europeu, e que se traduz, por exemplo, em alguns casos, como: *o' réis* (os réis) e *dé' reis* (dez reis)⁵, o que constituiria, segundo eles, em argumentos sólidos para a explicação da variação nos mecanismos de concordância nominal no português popular do Brasil.

Nessa mesma direção, transitou Silva Neto (1963) argumentando que as línguas românicas já possuíam tendência à simplificação, assim como ocorre no português popular. O contato linguístico estabelecido pelos africanos no processo de aquisição da língua portuguesa no Brasil contribuiu apenas, segundo o autor, para acelerar tal tendência.

Encontramos uma ideia similar em Câmara Jr. (1975) que não atribui importância alguma às línguas crioulas supostamente faladas no país pelos africanos, limitando-se apenas a reconhecer o papel exercido por elas no sentido de acelerar as mudanças já prefiguradas no próprio sistema da língua portuguesa. Mas é importante destacar também que o autor reconhece que as línguas africanas trazidas pelos escravos acabaram interferindo, sobretudo, nas variedades populares e na própria constituição da realidade linguística do país.

⁵ Aqui, um questionamento é inevitável: se a redução fonética da fricativa alveolar dental final teve consequências apenas no -s morfema plural, como explicar a mudança do paradigma flexional de verbo? Parece que não pode ter sido por redução fonética. Daí, a necessidade de um maior aprofundamento em torno dessa discussão.

Em suma, como se pode ver, essa discussão reflete os esforços de muitos estudiosos do português do Brasil em explicar a origem da versão popular da língua. Mas a deriva linguística, enquanto teoria, não foi a única perspectiva adotada para tal fim. Na subseção seguinte, discutiremos uma outra teoria explicativa que tem se mostrado bastante eficaz, sobretudo por embasar não só em dados empíricos, mas também na própria formação sócio-histórica do país.

1.1.2.2 Transmissão linguística irregular

Outra orientação relacionada ao tema da origem do português popular tem a ver com o conceito de transmissão linguística irregular proposta por Lucchesi e que tem ocupado um papel importante na agenda da pesquisa sociolinguística brasileira desde o início da década de 1990, graças aos estudos empreendidos por ele, que tem partido sempre de uma perspectiva da polarização sociolinguística no Brasil. O autor tem feito uma extensa pesquisa empírica, de campo em comunidades quilombolas do interior da Bahia, comunidades rurais cujos resultados encontram-se publicados no livro *Português afro-brasileiro* que teve um impacto muito grande na pesquisa linguística, chegando, inclusive, a mudar a orientação de uma visão de evolução interna da língua para uma agenda de pesquisa mais ampla que integra como questão central o contato linguístico como elemento desencadeador dos processos de variação que atingem, por exemplo, as regras de concordância nominal.

O conceito de transmissão linguística irregular é bastante explorado nos debates relacionados ao contato prolongado e massivo entre línguas. A sua definição está atrelada, visceralmente, a essa situação de contato responsável por emergir ou uma língua nova como *pidgin*⁶, crioulo, por exemplo, ou até mesmo “uma nova variedade histórica da língua que predomina na situação de contato” (LUCCHESI, 2000, p. 99). No caso do português, o que se verificou no Brasil foi um forte e pujante processo simultâneo de transmissão e aquisição irregular da língua do colonizador por parte dos povos dominados: escravos e indígenas. Esses povos apresentavam entre si uma enorme diversidade linguística e cultural revelando um plurilinguismo extraordinário. O colonizador foi implacável em impor seu senhorio aos

⁶ O conceito de língua pidgin pode ser encontrado em Baxter (1996), segundo o qual trata-se de um sistema linguístico um tanto reduzido cujo surgimento se dá em função da obrigatoriedade em que os falantes de línguas diversas se veem de estabelecer comunicação em algumas situações específicas, tais como: escravidão e comércio. Silva (2005) também discute o conceito dessa palavra, definindo-a como “uma língua simplificada cuja origem está ligada às situações de contato entre grupos de diferentes línguas entre si ininteligíveis, cujos falantes, por necessidade, têm de desenvolver atividades numa mesma situação geográfica e social”.

cativos, o que implicou a estes, no plano linguístico, uma forçosa necessidade de adquirir uma segunda língua, no caso, o português.

Ocorre que essa aquisição acabou se dando de forma um tanto precária por razões circunstanciais nas quais o contato dos escravos, muitas vezes, se restringia aos senhores e funcionários das fazendas. A esse respeito, Lucchesi (2001) é bastante enfático ao afirmar que:

[...] o português aprendido em situações sociais extremamente precária, e que era a língua de intercuro entre escravos e capatazes e senhores, e entre escravos de etnias diversas, constituiu o principal modelo para a nativização do português entre os descendentes desses escravos, sejam os provenientes de cruzamento de escravos de diferentes etnias, sejam os oriundos do cruzamento do colonizador branco com as mulheres escravas” (LUCCHESI, 2001, p. 101).

Como consequência dessa situação de contato fragmentário, emerge uma variedade de língua que se caracteriza, sobretudo, por apresentar uma estrutura gramatical tanto reduzida quanto simplificada, limitando-se a manter um sistema sintático capaz de atender as necessidades básicas e imediatas de comunicação. De acordo com Lucchesi (2003), algumas situações favoreceriam a tendência à redução e à simplificação das formas gramaticais, a saber:

- (i) o difícil acesso dos falantes das outras línguas aos modelos da língua alvo, sobretudo nas situações em que os falantes dessa língua alvo são numericamente muito inferiores aos falantes das outras línguas;
- (ii) o fato de os falantes dessas outras línguas serem, em sua maioria, adultos, não havendo, pois, o acesso aos dispositivos inatos que atuam naturalmente no processo de aquisição da língua materna;
- (iii) a ausência de uma ação normatizadora, ou seja, de uma norma ideal que oriente e restrinja o processo de aquisição/nativização, já que esse processo tem como objetivo o de fundamentalmente [ser] a comunicação emergencial com os falantes da língua alvo (LUCCHESI, 2003).

Pelo exposto, consideramos que o trabalho de Naro e Scherre (2007) teve um papel muito importante na pesquisa linguística do Brasil, mas essa posição que eles destacam, particularmente, no livro *Origem do português brasileiro* parece ter cada vez menos espaço na pesquisa linguística no Brasil. Até porque, as evidências empíricas têm demonstrado que em primeiro lugar não há nenhum resultado empírico hoje que aponte para uma deriva, uma mudança gradual e lenta de perda. Por exemplo, na norma culta, não se tem observado esse processo de perda. Aparentemente, pode existir, mas, na prática, é uma variação estável, uma

gradação geracional. E os trabalhos de análise tanto na concordância nominal quanto verbal da norma popular mostram um aumento da concordância e não a perda.

É interessante notar também que o grande argumento que Naro e Scherre (2007) apresentam para defender sua proposta era que a variação existia também em Portugal. E os estudos mais recentes têm derrubado também esse último argumento da hipótese, mostrando que a variação na concordância nominal e verbal em Portugal é praticamente inexistente. É um caso de flutuação de fala muito restrito a determinado contexto. Não tem nada a ver com o fenômeno massivo que se vê no Brasil cujas evidências apontam para um processo chamado de transmissão linguística irregular que é desencadeado em situações de contato. Nesse processo, é possível detectar, conforme Lucchesi (2009, p. 124), uma situação mais radical, tendo como resultado uma língua crioula que elimina, por exemplo, os elementos gramaticais da concordância nominal de número. Por outro lado, há uma situação de transmissão linguística irregular de tipo leve, da mesma natureza que aconteceu no Brasil por conta dos quatro milhões de africanos que compunham a base de nossa pirâmide social. Nessa situação, o que se tem é um quadro de variação e não de eliminação. Na essência, o processo é o mesmo, tem a ver com aquisição defectiva, imperfeita da língua do grupo dominante pela maioria de falantes das outras línguas: africanos e índios. Esse modelo imperfeito vai se tornando a língua materna dos descendentes. Foi o que aconteceu no Brasil por conta da opressão dos escravos que eram obrigados a deixar as suas línguas nativas e os chamados crioulos, que eram os que nasciam no Brasil. Eles não tinham acesso às línguas maternas dos pais e iam adquirindo o português nas senzalas de uma forma precária, daí esse processo de simplificação morfológica.

Portanto, a perspectiva da transmissão linguística irregular parece ter muito mais fundamentação empírica. A deriva, por sua vez, parece não ter nos tempos atuais muita consistência nem muita aceitação nos marcos atuais do desenvolvimento da teoria linguística. Até porque, não dá para pensar em tendências unilaterais de longo prazo, visto que qualquer discussão que advogue a favor de que as línguas tendem a simplificação se esbarra sempre na insuficiência de evidências empíricas. Ademais, é preciso considerar que, no geral, as línguas têm um mesmo nível de complexidade estrutural. Nesse sentido, ao conceber a existência de uma tendência geral nas línguas para simplificação, torna-se inevitável o reconhecimento de que há línguas gramaticalmente mais complexas e línguas menos complexas, o que, a nosso ver, parece um tanto contra intuitivo.

1.2 O PORTUGUÊS CULTO DO BRASIL

Ao assumirmos a nossa realidade linguística como heterogênea, reconhecemos, igualmente, na sociedade brasileira, a existência de uma realidade polarizada: português popular e português culto, cujos parâmetros de variação remetem-nos não só ao lugar de que procedem os falantes, mas a classe social e cultural em que eles se encontram. No caso do português culto brasileiro, alvo de nossa atenção nesse momento, o seu surgimento - muito mais recente se comparado com o português popular - está relacionado, segundo Castilho, diretamente com:

[...] a urbanização, que diferenciou a sociedade brasileira em dois níveis sociolinguísticos: a dos escolarizados e a dos analfabetos. As cidades trouxeram as escolas, os teatros, os livros e os jornais. Inicialmente imitando o português culto europeu (a classe administrativa brasileira era formada em Coimbra), pouco a pouco essa classe social encontrou sua personalidade, recolhendo formas populares mudando outras, até que a partir de 1920 se ergue a consciência de uma identidade linguística brasileira própria⁷.

O processo de gestação do português culto brasileiro estaria, segundo Lucchesi (2004, p. 77), vinculado aos padrões linguísticos europeus. As pequenas elites que compunham o quadro da administração portuguesa nos poucos centros urbanos do Brasil colonial se mostravam, conforme o autor, bastante zelosas e fiéis aos modelos de uso da língua oriundos da metrópole. Essa perspectiva sugere-nos a inegável constatação de que o português culto contemporâneo, falado e escrito pelas pessoas escolarizadas, seria legatário de um antecedente histórico que, nas palavras de Mattos e Silva (2008, p. 17), “seria próprio aos segmentos mais altos da sociedade colonial e teria o português europeu do colonizador como modelo a ser seguido”. A autora destaca ainda o fato de que

O português brasileiro culto começaria a definir-se, na segunda metade do século XVIII, quando, por força das diretrizes político-culturais pombalinas, a língua portuguesa passou a ser objeto de ensino sistemático no Brasil, embora precário, mas obrigatório, para fazer recuar a prevalência do ensino jesuítico, centrado na catequese, que priorizava a língua geral, e também no ensino da língua latina (MATTOS E SILVA, 2008, p. 16-17).

Esse português, fortemente vinculado à tradição escrita, manteve, em terras brasileiras, uma relação visceral com o português europeu, sendo falado pelos portugueses e luso-descendentes que sintetizavam a etnia branca do Brasil colônia. Dados levantados por

⁷ Texto extraído do Museu da língua portuguesa. Disponível em: www.estacaodaluz.org.br

Mattos e Silva (2004) dão conta de que essa população correspondia a 30% no país até a primeira metade do século XIX. Os 70% correspondentes ao restante da população, a maioria dos habitantes, portanto, eram, sobretudo, africanos e afro-descendentes, além de uma minoria de indígenas remanescentes do grande extermínio perpetrado contra eles por conta da sanha do conquistador. Esse cenário refletiu, do ponto de vista sociocultural, na configuração de uma sociedade marcada pelo signo do analfabetismo que se abatia entre 85% da população no final do século XIX, segundo estudos realizados por Ribeiro (1999). Essa realidade começou lentamente a mudar no primeiro período republicano quando o ensino laico passou a fazer parte da agenda política de então. Um indício dessa mudança lenta e gradual é o fato de que até a segunda década do século XX, 25% da população eram, conforme a autora, falantes do português culto brasileiro.

A quase ausência de escolarização no Brasil durante os períodos colonial e imperial - ressaltando o ensino jesuítico voltado à catequese e à formação de uma minoria - contribuiu para que o português culto falado na época continuasse, por um bom tempo, muito semelhante ao português culto europeu. Castilho chega a afirmar que “no século XIX, o português culto era imitação exata do português culto europeu”⁸. Foi preciso esperar a semana de arte moderna⁹ de 1922 para que o português culto brasileiro adquirisse uma caracterização muito mais abasileirada. Enquanto marco histórico em termos de ruptura com os padrões artísticos europeus, os ideais modernistas repercutiram, igualmente, na própria linguagem. O português brasileiro, que no romantismo¹⁰ já encontrara um ferrenho defensor de sua autonomia em relação ao português de Portugal, como é o caso de José de Alencar¹¹, encontra, no Modernismo, um ambiente bastante propício no qual “as letras nacionais passaram por um

⁸ Essa afirmação do autor consta numa entrevista dada por ele a agência Fundo de Amparo à Pesquisa do Estado de São Paulo-FAPESP acerca das origens do português brasileiro. Disponível em: <http://fronteirasnotempo.blogspot.com.br/2011/04/as-origens-do-portugues-brasileiro-pt.html>.

⁹ Considerada como o ápice do movimento modernista no Brasil, englobando as artes literárias, plásticas e musicais, a semana de arte moderna, que ocorreu entre os dias 13, 15 e 17 de fevereiro de 1922, no Teatro Municipal de São Paulo, tinha um propósito bastante definido: romper com as tradições estéticas europeias, no que tange as mais diversas manifestações artísticas, valorizando, sobretudo, a liberdade de criação. Houve, de fato, uma grande agitação cultural no meio letrado de modo que mesmo diante da resistência de muitos intelectuais da época, verificou-se, no plano linguístico, que muitos estudiosos passaram a assumir com mais espontaneidade e naturalidade a forma como os grandes escritores modernistas do Brasil escreviam.

¹⁰ O Romantismo no Brasil coincidiu com a independência política do país em 1822. O sentimento de nacionalismo aflorou tanto entre os setores mais influentes do cenário político quanto entre os intelectuais e escritores da época que já defendiam a caracterização do português como língua brasileira. No fundo, havia a sensação de que a independência em relação a Portugal só se completaria em sua plenitude quando o português falado no Brasil seguisse seu curso próprio sem, necessariamente, se manter submisso à sintaxe portuguesa.

¹¹ Em sua obra *Iracema*, cuja repercussão foi muito grande tanto no Brasil quanto em Portugal, José de Alencar assumiu de forma bastante precisa, no plano linguístico, uma linguagem muito mais original e autêntica, do ponto de vista dos padrões de uso por aqui adotados. Um exemplo é a colocação dos pronomes átonos cuja diferença em relação à forma usada em Portugal serviu de combustível para as ferozes críticas ao escritor.

grande impacto renovador, e onde figurava, pelo menos como ponto de pauta, a questão linguística” (FARIAS, 1997, p. 1015).

É importante, no entanto, evitar a temerária suposição de que dos ideais defendidos pelos modernistas tenham emergido a língua portuguesa culta do Brasil. Até porque, também entre os poetas modernos o que se advogava era a livre expressão literária que deveria se adequar à realidade brasileira. O que incluía, obviamente, o uso da língua falada no país. Daí, o uso inovador da linguagem como na poesia de Carlos Drummond de Andrade e na prosa de Guimarães Rosa, cuja característica principal é o regionalismo.

Por outro lado, a inserção de uma temática brasileira no contexto da produção literária moderna demandava a adoção de um comportamento linguístico que pudesse traduzir com mais fidelidade a realidade do país. Ora, considerando a autoridade intelectual dos escritores modernistas cuja farta produção literária pôs em evidência peculiaridades linguísticas brasileiras que, até então, eram ignoradas nos círculos linguísticos mais europeizados, não é de se espantar que o português culto brasileiro falado pelas pessoas mais escolarizadas tenha recebido forte influência dos modernistas. Castilho chega a afirmar que

[...] os modernistas tiveram um papel muito importante, sobretudo Mário de Andrade. Ele criou biblioteca, departamento de cultura, fundou a revista do Arquivo Municipal. Houve uma grande agitação cultural e as pessoas foram assumindo com mais naturalidade a forma como elas escreviam¹².

Vale a pena destacar também que o processo de urbanização mais intenso a partir de meados do século XX desencadeou uma onda migratória no Brasil com multidões de indivíduos oriundos dos mais variados recantos do país convergindo para as grandes cidades em busca de melhores condições de vida. O quadro linguístico que se compôs em função de tal convergência foi o de um autêntico laboratório dialetal em que as variedades diatópicas refletiam a diversidade sociocultural dos brasileiros. Contudo, essa mobilidade interna trouxe sérios problemas para a organização do espaço urbano, principalmente nos grandes centros. As cidades não estavam preparadas, muito menos planejadas urbanisticamente para acolher um contingente populacional imenso. E o resultado foi a aglomeração desordenada nas regiões mais periféricas, intensificando ainda mais os graves problemas sociais. E é justamente aí que entram em cena as variedades diastráticas, denunciando a estrutura social dos falantes.

¹² Conferir texto completo em: www.estacaodaluz.org.br.

Esse cenário aqui delineado, sinteticamente, pôs em evidência, dentre as várias variedades dialetais, aquela variedade de prestígio falada pelos utentes escolarizados, inseridos, portanto, no contexto da cultura letrada. Essa variedade da língua portuguesa, também conhecida como português culto brasileiro se aproxima da norma padrão estabelecida pela gramática tradicional, porém, não se constitui como essa em “um constructo sociocultural, uma norma no sentido mais jurídico do termo, uma espécie de lei linguística que prevê a condenação e a punição dos infratores” (BAGNO, 2007, p. 98). A melhor definição para ela é aquela que a relaciona aos níveis mais alto de escolarização.

De fato, a inserção na cultura letrada tem sido um critério fundamental para se classificar o informante como falante do português culto. O projeto NURC (Norma Urbana Culta) estabeleceu o nível superior de escolaridade e o contexto urbano na pesquisa realizada em cinco capitais brasileiras com mais de um milhão de habitantes na época: Recife, Salvador, Rio de Janeiro, São Paulo e Porto Alegre, com objetivo bastante claro: descrever e documentar a norma do português culto falado no Brasil. A coleta dos materiais gravados ocorreu entre 1970 e 1977. Os dados coletados serviram de base para diversas pesquisas sociolinguísticas acerca dos fenômenos de variação nessa modalidade da língua. Além disso, o pioneirismo do projeto NURC abriu um precedente para que outras pesquisas sobre o português culto fosse empreendida no país, como é o caso do Projeto de Gramática do Português Falado, com ênfase na Gramática do Português Culto Falado no Brasil que redundou, segundo Castilho (2007, p. 100) em oito volumes na fase inicial, entre os anos de 1988 e 2002, que serviram de base para publicação de vários ensaios. Essa farta produção acabou sendo condensada, na segunda fase do projeto, em 2004, em cinco volumes: *Construção do texto*, *Classes de palavras e construções gramaticais*, *Construção da sentença*, *Construção morfológica da palavra* e *Construção fonológica da palavra*.

Reproduzimos logo abaixo um quadro comparativo apresentado por Castilho no Museu da língua portuguesa no qual estão perfeitamente descritas as características do português culto ao lado daquelas relacionadas ao português popular:

Quadro 1: Características do português culto

Português brasileiro popular	Português brasileiro culto
PRONÚNCIA DAS VOGAIS E DOS DITONGOS	
Ditongação das tônicas seguidas de sibilante no final das palavras: <i>mêis, luiz</i>	Essas vogais são preservadas: <i>mês, luz</i> .
Átonas iniciais podem nasalar-se: <i>enzame, inducação, inleição</i> .	Mantém-se a átona inicial, flutuando sua pronúncia como <i>exame / izame, educação / iducação</i>

Abertura das átonas pretônicas no Nordeste (<i>còvardi, nòturno, nèblina, rècruta</i>), fechamento no Sul (<i>covardi, noturno, etc.</i>). Fechamento maior em palavras dissilábicas, donde <i>filiz, chuver</i>	Mesmos fenômenos.
Queda das vogais átonas postônicas nas proparoxítonas: <i>pêzgu, cosca, oclos</i> , por <i>pêssego, cócegas, óculos</i> . Com isso, predominam as paroxítonas.	Mantêm-se as átonas postônicas nas proparoxítonas
Vogais átonas finais <i>-e, -o</i> são mantidas em algumas regiões, e fechadas em outras encontrando-se as pronúncias <i>pente – penti, lobo – lobu</i> .	Mesmos fenômenos.
Perda da distinção entre ditongos e vogais em contexto palatal: monotongação em <i>caxa, pexe, bejo, quejo</i> ; ditongação em <i>bandeija, feichar</i>	Não ocorre a ditongação
Desnasalação e monotongação dos ditongos nasais finais: <i>hómi, faláru</i> .	Os ditongos nasais são mantidos: <i>homem [òm~ey], falaram [falárãw]</i> .
Monotongação dos ditongos crescentes átonos em posição final: <i>ciença, experiença, negoço</i> .	Manutenção desses ditongos: <i>ciência, experiência, negócio</i> .
PRONÚNCIA DAS CONSOANTES	
Retroflexão do <i>r</i> na área dos falares caipiras, seja no final ou na posição inicial de sílaba e nos grupos consonantais: <i>porta, caro, cobra</i> . No Nordeste e no Rio de Janeiro, vibração posterior. No Sudeste e Sul, vibração anterior.	Mesmos fenômenos, com a tendência a discriminar o <i>r</i> retroflexo em situações formais.
Troca de <i>l</i> por <i>r</i> em final de sílaba e em grupos consonantais: <i>marvado, pranta</i> .	Manutenção do <i>l</i> : <i>malvado, planta</i> .
Troca de <i>v</i> por <i>b</i> em palavras tais como <i>barrer, bassoura, berruga, bespa</i> , em Pernambuco, Bahia e São Paulo.	Manutenção de <i>v</i> : <i>varrer, varrouira, verruga, vespa</i> .
As dentais <i>t</i> e <i>d</i> em posição final (1) podem ser mantidas como tais, (2) palatizadas, como em <i>denti, pòdi</i> , (3) africadas como em <i>dentʃi, pòdʒi</i> .	Mesmos fenômenos.
Iodização da palatal <i>lh</i> : <i>oreya, vèy</i> .	Manutenção da palatal: <i>orelha, velho</i> .
Espiração e perda de <i>-s</i> final: <i>vamoh > vamo; poi</i>	Manutenção da sibilante: <i>vamos, pôs. > pô</i> .
MORFOLOGIA	
Morfologia nominal e pronominal	
Perda progressiva do <i>-s</i> para marcar o plural, que passa a ser expresso pelo artigo: <i>os hòmi, as pessoa</i> .	Manutenção das regras redundantes de marcação do plural, salvo na fala rápida: <i>os homens, as pessoas</i> .
Perda do valor do sufixo <i>-ior</i> nos comparativos de superioridade, utilizando-se o advérbio <i>mais</i> : <i>maismiô, mais piô</i> .	Preservação do valor comparativo do sufixo <i>-ior</i> : <i>melhor, pior</i> .
Alterações no quadro dos pronomes pessoais: generalização do reflexivo <i>se</i> para a primeira pessoa (<i>eu se esqueci, nós não se falemo mais</i>), perda do pronome <i>o</i> , generalização do pronome <i>lhe</i> , substituição de <i>tu</i> por <i>você</i> no centro do país, substituição de <i>nós</i> por <i>gente</i> .	O pronome reflexivo ou mantém sua pessoa gramatical, na terceira pessoa (<i>ele se esqueceu</i>) ou é omitido (<i>eu esqueci</i>). A perda de <i>o</i> na língua falada se difunde, mantendo-se apenas na língua escrita. Usa-se <i>tu</i> apenas nas regiões Norte e Sul do país, neste caso sem com ele concordar o verbo: <i>tu sabede uma coisa?</i>
Redução do quadro dos pronomes possessivos para <i>meu / seu / dele</i> , com perda progressiva de <i>teu</i> nas regiões em que desapareceu <i>tu</i> .	Mesmas características. O pronome <i>teu</i> pode aparecer em contextos marcados, alternando com <i>seu</i> : <i>Meta-se com os seus negócios, isto não é datua conta!</i>
Redução dos pronomes demonstrativos a dois tipos: <i>este/esse</i> , para indiciar objetos próximos ou para retomar informações próximas, mantendo-se <i>aquele</i> para indiciar objetos e informações remotas.	Mesmas características.

Generalização do pronome relativo <i>que</i> , perdendo-se <i>cujo, onde</i> .	Mesmas características.
Morfologia verbal	
Elevação da vogal temática no pretérito perfeito do indicativo: <i>fiqueмо, faleмо, bebimu</i> .	Manutenção da vogal temática, continuando indistintos o presente e o pretérito: <i>ficamos, falamos, bebemos</i> .
Simplificação na morfologia de pessoa, dadas as alterações no quadro dos pronomes pessoais, reduzindo-se a conjugação a apenas duas formas diferentes: <i>eu falo, você / ele / a gente / eles fala</i> . Por hipercorreção, pode-se ouvir <i>a gente falamos</i> .	A morfologia de pessoa reduz-se a três, às vezes a quatro formas diferentes: <i>eu falo, você / ele / agente fala / eles falam</i> .
SINTAXE	
Simplificação da concordância nominal, expressa apenas pelo determinante (como em <i>as pessoa</i>), e acentuada quando o substantivo e o adjetivo vêm no diminutivo (<i>aqueles cabelim branquim</i>). A concordância é ainda visível quando há saliência fônica diferenciando a forma singular da forma plural, como em <i>as colheres</i> .	Manutenção da concordância nominal com redundância de marcas: <i>as pessoas, aquelescabelinhos branquinhos</i> .
Simplificação da concordância do verbo com o sujeito: <i>as pessoa fala, fala, mas não resolve nada</i> . Ocorrendo saliência fônica entre as pessoas do verbo, mantém-se a concordância: <i>as pessoa saíru, mas elas são bão</i> .	Mantém-se a concordância do verbo com o sujeito, mas a regra pode não se aplicar quando o sujeito é posposto e separado do verbo por expressões várias: <i>Faltou mesmo depois de tanta luta as respostas mais interessantes</i> .
Predominância do sujeito exposto e colocado antes do verbo, evitando-se o sujeito posposto;	Mesma característica. Sujeito “pesado”, isto é, constituído por muitas sílabas tende a pospor-se, mas a sintaxe torna-se progressivamente mais rígida.
Objeto direto pronominal exposto pelo pronome <i>ele</i> (<i>eu vi ele</i>) ou por <i>lhe</i> (<i>eu não lhe conheço</i>). Objeto indireto exposto por pronome demonstrativo neutro e complemento oblíquo tendem a aparecer antes do verbo: <i>Isso eu quero, Isso eu preciso</i> .	Discreta preferência pelo objeto direto omitido: <i>eu vi Ø</i> . Na fala culta espontânea é comum dizer-se <i>eu vi ele</i> , mas ainda é raro o uso de <i>lhe</i> como objetodireto. Mesmas características nos demais casos.
Abundância de construções de tópico com retomada pronominal no interior da oração: <i>A menina, elachevou agora mesmo</i> .	Mesma característica.
Preferência pela oração relativa cortadora, em que se omite a preposição antes do pronome relativo (<i>perdi a revista que a capa estava rasgada</i>) e pela relativa copiadora, em que se insere pronome pessoal depois do relativo (<i>o menino que ele chegou trouxe a correspondência</i>). Nos dois casos, nota-se que o relativo se “despronominaliza” e é cada vez mais apenas uma junção.	Preferência pela oração relativa padrão, sobretudo na variedade escrita: <i>perdi a revista cuja capa estava rasgada, o menino que chegou trouxe a correspondência</i> . Na variedade falada espontânea já se encontram as relativas cortadora e copiadora.
Preferência pela oração substantiva “dequeísta”: <i>Ele falou de que não sabia de nada</i> .	Preferência pela oração substantiva “não-dequeísta”: <i>Ele falou que não sabia de nada</i> .

Fonte: Museu da Língua Portuguesa (estacaodaluz.org.br).

Essas características¹³ observadas sugerem-nos que, em muitas ocorrências envolvendo o português culto, é possível encontrar traços linguísticos graduais, ou seja,

¹³ Castilho sintetiza as características de ambas as variedades, afirmando que “(1) não há uma oposição categórica entre fala popular e fala culta, ocorrendo em muitos casos um compartilhamento de propriedades; (2) em certos casos a preferência culta exclui fortemente a preferência popular; (3) em situações informais, diminui a distância entre essas variedades, e o falante culto pode aproximar-se bastante da execução popular; ainda que não em todos os casos; (4) as variedades populares flutuam de acordo com a região geográfica, mas a fala culta é um pouco mais homogênea, sobretudo em sua forma escrita”.

fenômenos linguísticos que “aparecem na fala de todos os brasileiros, independentemente de sua origem social, regional etc.” (BAGNO, 2007, p. 142). Essa tendência, na norma culta, para assimilação de algumas formas não previstas na norma-padrão pode ser um forte indício de que é preciso urgentemente repensar os padrões de uso normativo da língua portuguesa no Brasil, tendo em vista a imensidão do país tanto geograficamente quanto demograficamente, se comparado com Portugal de onde emanam as prescrições normativas. As mesmas que o gramático português Cândido de Figueiredo, no início do século XX, preconizava na obra *O que se não deve dizer. Bosquejos e notas de filologia portuguesa ou consultório popular de enfermidades da linguagem*.

Os esforços de muitos linguistas empenhados em pesquisar a realidade linguística brasileira, notadamente, no que se refere ao português culto, têm redundado em inúmeros projetos e trabalhos acadêmicos nos quais os padrões de uso concreto da língua têm sido mapeados, com base em dados empíricos, mostrando que a língua portuguesa brasileira, falada pelos segmentos mais escolarizados do país, parece ganhar estatuto próprio em termos de autonomia em relação ao português europeu, sobretudo, no âmbito da morfossintaxe.

Foi com base nesses estudos relacionados ao processo de formação e desenvolvimento dessas duas vertentes do português brasileiro: a popular e a culta, que começou, a partir da década de 1990, a se postular, para o contexto linguístico do Brasil, a configuração de uma realidade sociolinguística polarizada. Na seção a seguir, é possível entender melhor os bastidores desse debate.

1.3 A POLARIZAÇÃO SOCIOLINGÜÍSTICA BRASILEIRA

Para caracterizar a realidade linguística do Brasil, Lucchesi (1994, 1998, 2001, 2002 e 2006) tem partido sempre da perspectiva de uma polarização da história sociolinguística do país¹⁴. A concepção polarizada adotada pelo autor se baseia no fato de que

desde o início da colonização até a Proclamação da República, enquanto uma reduzida elite concentrada nos incipientes centros urbanos guardava uma profunda fidelidade aos modelos de uso da língua provenientes de Portugal, nas imensidões do interior do Brasil grandes contingentes de índios aculturados e africanos

¹⁴ O autor vem se debruçando desde a década de 1990 sobre a formulação de uma análise da realidade sociolinguística brasileira, definindo-a como uma realidade sociolinguística polarizada. Com isso, ele tem procurado lançar as bases para a análise sócio-histórica da mudança do português brasileiro, chegando, inclusive, a derrubar certos equívocos que orientavam a pesquisa no Brasil, por exemplo, a ideia de que existe uma tendência gradual para a perda das regras de concordância nominal e verbal. Em seus estudos, o autor tem demonstrado que todas as pesquisas empíricas que têm sido feitas nas últimas décadas sobre o português popular brasileiro demonstram exatamente o contrário.

escravizados adquiriam a língua portuguesa em condições as mais precárias; e essa língua segunda defectiva se ia convertendo em modelo para a nativização do português entre os descendentes mestiços e endógamos desses segmentos, desencadeando profundas alterações na gramática da língua portuguesa assim adquirida, socializada e nativizada. (LUCCHESI, 2006, p. 87-88).

O cenário linguístico que se estabeleceu como reflexo dessa realidade sócio-histórica é descrita por ele da seguinte forma:

Configura-se, assim, o processo histórico de formação das duas grandes normas do português brasileiro: a *norma culta*, derivada do uso linguístico de uma elite escolarizada, e a *norma popular*, que emerge do uso da grande maioria da população do país, desprovida de educação formal e dos demais direitos da cidadania, com os previsíveis reflexos na língua da pluralidade étnica que está na base da sociedade brasileira. (LUCCHESI, 2006, p. 88).

O que se tem notado tanto nos estudos de Lucchesi quanto nos de Matose Silva ¹⁵ (2004) é que essa polarização sociolinguística do português brasileiro começa a se estruturar já a partir de meados do século XVI, período em que é inaugurado, de fato, a colonização do Brasil. Portanto, o contexto colonial que perdurou até o início do século XIX apresentou alguns cenários históricos e sociais que se constituíram como eventos fundamentais para se entender tanto as origens sócio-históricas da polarização linguística aqui discutida, como também o estabelecimento do seu padrão normativo. A fim de melhor compreender os eventos que ensejaram o desenvolvimento dessa realidade polarizada da língua portuguesa em solo brasileiro, remetemo-nos, ainda que de forma breve, a alguns fatos importantes da nossa historiografia, incluindo, obviamente, acontecimentos subsequentes ao Brasil colônia.

Em relação aos três primeiros séculos de afirmação da nossa condição de país colonizado, é importante destacar o fato de que o Brasil se caracterizava, sobretudo, por ser o que Cunha (1985, p. 17) chama de “vasto país rural”. O autor destaca ainda a inexistência de importantes centros culturais nas cidades e vilas, fato esse que refletiu na própria organização social do interior do país onde se encontrava a imensa maioria da população. Essa situação é vista por Lucchesi como bastante antagônica do ponto de vista linguístico. Conforme o autor

Nesse universo, temos, de um lado, os pequenos centros urbanos, onde se situavam os órgãos da administração colonial, sob forte influência cultural e linguística da metrópole. A elite colonial era naturalmente bastante zelosa dos valores europeus,

¹⁵ Na verdade, os autores reconhecem também uma divisão tripartida da realidade linguística brasileira, no contexto atual, com a seguinte configuração: norma-padrão, norma culta e norma popular. Para um estudo mais aprofundado do tema, vale a pena consultar as seguintes obras dos autores: LUCCHESI, Dante. **Norma linguística e realidade social**. In: BAGNO, Marcos (org). *Linguística da norma*. São Paulo: Loyola, 2004. / MATTOS E SILVA, Rosa Virgínia. **O português são dois: novas fronteiras, velhos problemas**. São Paulo: Parábola Editorial, 2004.

buscando assimilar e preservar ao máximo (o que é previsível nessas situações) os modelos de cultura e de língua vindos d'além-mar (LUCCHESI, 2004, p. 77).

Estaria aí na base desse vínculo com os padrões linguísticos europeus o que o autor chama de “antecedente histórico da variedade culta do português brasileiro” (cf. p. 77).

Por outro lado, o português foi se difundindo de forma cada vez mais pujante entre as populações interioranas, refletindo toda uma diversidade sociocultural peculiar a esses povos: africanos, afrodescendentes e indígenas. Em sua percepção refinada da sócio-história brasileira, sobretudo em sua dimensão linguística, Lucchesi afirma, peremptoriamente, que

a outra vertente da formação da língua no Brasil fincou raízes no interior do país, para onde se dirigiu a maior parte da população no período colonial. Fora dos reduzidos centros da elite, nas mais diversas regiões do país, o português era levado, não pela fala de uma aristocracia de altos funcionários ou de ricos comerciantes, mas pela fala rude plebeia dos colonos pobres. Além disso, a língua portuguesa era adquirida nas situações as mais precárias pelos escravos, que muitas vezes preferiam se comunicar entre si usando uma língua franca africana. Sob essas ásperas condições, a língua portuguesa se foi disseminando entre a população pobre, de origem predominantemente indígena e africana, nos três primeiros séculos da história do Brasil (LUCCHESI, 2004, P. 77).

Essa variedade do português modificada em função das mudanças estruturais introduzidas pelo contato entre línguas foi chamada por Mattos e Silva (2004) de “português geral brasileiro”. Uma espécie de antecedente histórico do português popular da atualidade.

Com a chegada de milhares de imigrantes ao país no final do século XIX, oriundos tanto de países europeus quanto da Ásia, o Brasil se torna um celeiro linguístico ainda mais privilegiado. Essas pessoas vieram trabalhar nas lavouras, o que os mantinha na base da pirâmide social brasileira, como bem destacou Lucchesi. Ainda segundo o autor, “o modelo mais acessível de que dispunham para a aquisição do português era o proveniente dos capatazes e dos trabalhadores braçais locais que, em sua maioria, eram ex-escravos africanos e seus descendentes nativos e /ou mestiços” (LUCCHESI, 2004, p. 80). O português aprendido nessas circunstâncias era justamente o português popular. Mas o autor lembra que devido ao nível de instrução de grande parte desses imigrantes não foi difícil para eles alcançarem uma rápida ascensão social, passando, inclusive, a conviver linguisticamente com a classe média da época: os brancos e seus descendentes, graças a mão-de-obra especializada que os distinguia de longe dos demais trabalhadores brasileiros. O reflexo desse convívio foi a absorção por parte do português culto de certas estruturas da fala popular.

Vale ressaltar também que no século XX o Brasil se tornou um país capitalista. E o primeiro reflexo dessa condição política foi o início do processo de industrialização que, por

sua vez, acelerou a urbanização. Com isso, assiste-se, sobretudo a partir da década de 1940, a uma radical alteração no quadro demográfico brasileiro em função do êxodo rural. Um contingente inumerável de pessoas sai das localidades rurais e começam a abarrotar os grandes centros urbanos em busca de melhores condições de vida. Para Bortoni-Ricardo essa situação, aliada a difusão dos meios de comunicação de massa, instalou-se, no plano linguístico,

um processo de diglossia, onde atuam duas forças antagônicas: por um lado, o padrão tradicional de redução flexional da própria língua, exarcebado pela situação de contato entre dialetos diferentes; por outro, a pressão do prestígio da norma culta, imposta pela ação da escola, dos meios de comunicação e do *status* das classes mais favorecidas (BORTONI-RICARDO, 2005, P. 33).

O contexto sociolinguístico hodierno, no Brasil, ainda é fortemente marcado pela coexistência das normas popular e culta caracterizadoras da realidade linguística bipolar, embora Lucchesi (2001, p. 107) defenda uma certa atenuação ou diminuição em termos de distância entre ambas as normas em virtude da mudança no comportamento linguístico motivada pelo acesso da população das camadas mais baixas da sociedade à televisão, ao rádio, à escola e até mesmo aos meios de transportes mais modernos que proporcionam o contato direto entre as pessoas. Essa mudança observada pelo autor se daria em direção à aquisição de modelos da norma culta¹⁶.

Essa breve e sumária exposição acerca de um tema tão complexo e, ao mesmo tempo, tão instigante, como é o caso da polarização linguística, evidencia um outro aspecto relativo ao português brasileiro, como bem ilustra Lucchesi (1994, p. 27), a saber: o de que além de heterogêneo e variável, a língua portuguesa aqui transplantada tornou-se plural e polarizada pelas razões já discutidas logo acima. A consequência dessa realidade pode ser vislumbrada no cenário em que a língua se impõe enquanto veículo de comunicação e de difusão cultural cuja variabilidade pode se manifestar com maior ou menor intensidade segundo o contexto social em que se encontram inseridos os falantes. Um ambiente em que há um predomínio do uso da norma culta pode refletir uma maior exposição do falante à cultura

¹⁶ Em seu artigo *Norma linguística e realidade social*, que compõem uma antologia de textos organizados por Bagno (2004), Lucchesi discute essa tendência de mudança envolvendo as normas popular e culta. Ele chega a afirmar que já “no decorrer do século XX, enquanto no português popular se verifica uma tendência de mudança “para cima”, não em direção aos padrões normativos, mas em direção ao padrão urbano culto (ou semiculto), no português culto, assiste-se a uma tendência de mudança de afastamento do padrão normativo de matiz europeu, uma mudança que se pode definir como “para baixo”. Nessa mesma direção, Silva (2005, p. 80-81), citando o autor em questão, lembra que essas duas tendências observadas no português brasileiro “apontam para um processo de convergência” em que “as duas variedades de língua estariam apresentando diferenças em relação aos padrões de uso que subjazem à norma culta e à norma vernácula”.

letrada. Enquanto a norma popular emerge de forma bastante sistemática naqueles ambientes em que a maioria das pessoas são privadas dos direitos mais elementares da cidadania, como o acesso à educação formal, por exemplo. A superação do antagonismo que envolve os usos das normas popular e culta na sociedade brasileira¹⁷, bem como o pesado estigma social que se abate sobre a fala popular só serão efetivamente mitigados quando a realidade socioeconômica da grande maioria da população for muito menos desigual.

¹⁷ Os trabalhos de Lucchesi não deixam dúvida de que, historicamente, a polarização sociolinguística do Brasil está associada exatamente a essa oposição entre a fala das classes populares e a fala da elite letrada. Entre a elite letrada, por exemplo, é possível imaginar que as regras de concordância nominal tendem a variar muito pouco e o quadro identificável é um quadro de variação estável e não de mudança em curso. As diferenças do ponto de vista do comportamento linguístico entre as classes baixas e a elite brasileira, bem como a avaliação social da mudança criaram um fosso que separa a realidade sociolinguística brasileira. Esse abismo sociolinguístico reflete nada mais nada menos do que a nossa realidade socioeconômica desigual. E, historicamente, tal abismo ainda foi maior por conta da própria formação histórica da sociedade brasileira em que a base da nossa sociedade foi formada por escravos africanos e índios aculturados em sua grande maioria. E não é à toa que hoje na base da pirâmide social brasileira predominam os afrodescendentes ou os índios descendentes a depender da região. Esse processo maciço de contato entre línguas é que vai cavar historicamente esse fosso. Com a chamada revolução de 1930 de Getúlio Vargas se inicia no Brasil um processo de industrialização de uma sociedade até então rural e baseada no latifúndio. Esse processo vai desde então mitigando, atenuando essa polarização. Daí o porquê as classes populares estarem avançando em direção à norma culta, como bem tem ressaltado Lucchesi. Só que o caráter tardio e dependente do desenvolvimento do capitalismo no Brasil tem contribuído, consideravelmente, para a manutenção desse quadro de polarização sociolinguística, conforme verificamos na atualidade.

2 FUNDAMENTAÇÃO TEÓRICA

Para discutir as bases teóricas sob as quais nos apoiamos para a realização do presente estudo, julgamos pertinente dividir esse capítulo em duas seções. Assim, na primeira, apresentamos algumas considerações breves acerca dos pressupostos teóricos que nortearam nossa pesquisa. Na segunda seção, tratamos de alguns estudos já realizados sobre a variação na concordância nominal de número no SN, no português brasileiro.

2.1 ASOCIOLINGÜÍSTICA VARIACIONISTA

Nesta seção, nos propomos, preliminarmente, a discutir, de forma bastante resumida, o contexto dos estudos linguísticos que antecedeu o advento da sociolinguística laboviana. Em seguida, discutiremos a teoria da variação e mudança que serviu de base para o nosso estudo, sobretudo, no que tange à variante *concordância nominal de número no SN*, bem como, os princípios teóricos fundamentais da sociolinguística variacionista.

2.1.1 O advento da sociolinguística

Reportando-nos à época imediatamente anterior as ideias saussurianas sobre a linguagem, mais precisamente do século XVIII até o início do século XX, veremos que por quase duzentos anos, portanto, os estudos linguísticos primavam, sobretudo, por um posicionamento eminentemente comparativo em que eram comparados textos escritos em diferentes línguas a fim de encontrar similitudes entre ambas que reforçassem a tese de que tais línguas pertenciam a uma única família. No fundo, havia um esforço considerável em provar a existência de uma língua mãe que dera origem a todas as demais. O próprio texto do Curso de Linguística Geral, discutindo a história da linguística que se desdobrara, segundo o autor, em três períodos, afirma que “o terceiro período começou quando se descobriu que as línguas podiam ser comparadas entre si. Tal foi a origem da Filologia comparativa ou da Gramática comparada” (SAUSSURE, 2012, p. 32). A esse respeito, Benveniste (2005) destaca, igualmente, que “uma fase nova abre-se no início do século XIX com a descoberta do sânscrito. Descobre-se, ao mesmo tempo, que existe uma relação de parentesco entre as línguas a partir daí chamadas indo-européias” (BENVENISTE, 2005, p. 20-21).

Não há dúvida de que o sistema colonial europeu, movido por interesses colonialistas, acabou possibilitando que muitos estudiosos europeus tivessem contato mais

regular com a língua sânscrita. E o resultado foi justamente a comparação dessa língua com outras línguas como, por exemplo, o grego e o latim que eram línguas muito importantes na cultura europeia. Com isso, foi possível uma percepção mais nítida do quanto as línguas mudavam regularmente, o que permitiu o desenvolvimento de toda uma teoria das mudanças fonéticas, partindo do princípio básico de que tais mudanças eram sistemáticas e regulares. A linguística nessa época era considerada parte inerente tanto da história quanto da biologia, conforme se adotasse a perspectiva de Franz Bopp ou de Schleicher¹⁸. O fragmento abaixo ajuda a visualizar melhor esse contexto:

[...] a linguística consistia essencialmente numa genética das línguas. Fixava-se, para tentar estudar a evolução das formas linguísticas. Propunha-se como ciência histórica, e o seu objeto era, em toda parte e sempre, uma fase das histórias das línguas. (BENVENISTE, 2005, p. 21)

Mas a ciência da linguagem moderna vai atingir a sua maturidade no início do século XX, a partir da publicação póstuma¹⁹ do livro de linguística geral de Ferdinand de Saussure.

O autor fez uma revolução no estudo da linguagem. Com ele, a pergunta norteadora da pesquisa linguística não era mais como as línguas mudam, mas como funcionam.

Ele passou a estudar a língua sincronicamente e não mais através do tempo, ou seja, a língua num determinado momento do seu funcionamento. Ele, inclusive, sugeriu que o linguista deveria abstrair, ou seja, ignorar a história da língua já que o falante falava a língua sem saber nada da história dela. Pelo menos é o que se depreende do fragmento abaixo:

A primeira coisa que surpreende quando se estudam os fatos da língua é que, para o indivíduo falante, a sucessão deles no tempo não existe: ele se acha diante de um estado. Também o linguista que queira compreender esse estado deve fazer *tábula rasa* de tudo quanto produziu e ignorar a diacronia. Ele só pode penetrar na consciência dos indivíduos que falam suprimindo o passado. A intervenção da história apenas lhe falsearia o julgamento (SAUSSURE, 2012, p. 123).

Vale lembrar que na teoria linguística Saussuriana a língua é concebida como um sistema de oposições regido por sua lógica interna. E essa concepção de língua sincrônica, como um sistema organizado para ter sua função, passou a orientar a pesquisa linguística

¹⁸ Franz Bopp é um dos principais criadores da gramática comparada. Na obra *Sistema da conjugação do sânscrito, ele estudou as relações que unem o sânscrito ao germânico, ao grego, ao latim etc* (CLG, 2012, p. 32). August Schleicher defendeu uma perspectiva biológica da linguagem. Para ele, a língua é um organismo natural que evolui.

¹⁹ A obra foi publicada em 1916, três anos após a morte de Saussure, por seus alunos e discípulos Charles Bally e Albert Sechehaye. É fruto de um apanhado de três cursos que ele havia ministrado na universidade de Genebra. O Primeiro curso aconteceu de 16 de janeiro a 3 de julho de 1907, com seis alunos matriculados; o segundo, da primeira semana de novembro de 1908 a 24 de julho de 1909, com onze alunos matriculados; o terceiro curso, de 28 de outubro de 1910 a 4 de julho de 1911, com doze alunos matriculados (CLG, 2012, p. 14-15).

desde a década de 1920 até o final da década de 1950 e início da década de 1960. Foi a partir daí que surgiram dois outros modelos que orientaram a pesquisa linguística desde então até a atualidade: a gramática gerativa, liderada por Noam Chomsky e a Sociolinguística variacionista de William Labov. É esse segundo modelo programático que nos interessa na presente discussão.

2.1.2 Teoria da variação e mudança

A sociolinguística se constituiu em um modelo científico de pesquisa empírica, inaugurado na década de 1960 por William Labov. Trata-se de um programa de pesquisa que passou a assumir, teoricamente, que a mudança e a variação não comprometem o funcionamento da língua, ao contrário do que supunham os estruturalistas. Os pesquisadores que abraçaram essa nova perspectiva teórica começaram a estudar exatamente de que modo se dá a relação entre a variação observável em cada momento da língua e o processo de mudança que vai acontecendo de forma permanente e ininterrupta em todas as línguas humanas em uso. Como bem destaca Lucchesi (2012)

[...] a sociolinguística surge como uma resposta à incapacidade do formalismo linguístico em tratar da questão da mudança. Para construir o objeto de estudo da Linguística Moderna, Saussure retirou a língua do seu devir histórico. Definindo a análise estrutural como exclusivamente sincrônica, o modelo saussuriano tornou-se incapaz de lidar com a questão da mudança (LUCCHESI, 2012, p. 793).

Portanto, a teoria da variação e mudança linguística, também conhecida como teoria variacionista, representou um marco histórico em termos de estudos e pesquisas linguísticas no âmbito da ciência da linguagem. William Labov, conforme já mencionado, tornou-se o nome mais representativo da sociolinguística. Foi ele quem lançou, de fato, as bases tanto de natureza teórica quanto empírica para se estudar os fenômenos linguísticos²⁰. Para ele, assim

²⁰ O estudo clássico, o modelo da análise sociolinguística foi feito por Labov em 1963 na Ilha de Martha's Vineyard. Em seu livro *Padrões Sociolinguísticos* (cf. p. 19-62) o autor nos apresenta um relato desse estudo, acompanhado, inclusive, de uma caracterização social e espacial da ilha que, segundo ele, na época, uma ilha de pescadores pouco habitada, na costa leste dos Estados Unidos e que estava começando a se tornar um destino turístico famoso. Ele relata ainda que a ilha, em tal época, era muito pouco habitada e os veranistas estavam começando a invadir esse espaço que veio a se tornar um balneário de luxo. Em sua pesquisa, Labov analisou, entre outras coisas, o processo de elevação dos ditongos do inglês /ay/ e /aw/ e constatou que os falantes usavam mais as variantes [əɪ] e [əʊ] que eram uma característica dialetal da ilha. Na verdade, essa característica dialetal estava na contramão da mudança do inglês que ia, por exemplo, de [əɪ] para /ay/, que era a pronúncia padrão no inglês. Então ele tentou explicar porque estava havendo esta retomada da pronúncia arcaica naquela localidade. Com isso, ele vai dar um salto de qualidade em seu estudo, já que a dinâmica da análise sociolinguística era estabelecer as correlações do plano do encaixamento social, a distribuição da variação na estrutura social da comunidade, o que se traduzia, por exemplo, em análises envolvendo o comportamento dos homens e das

como para Weinreich e Herzog (2006), “será necessário aprender a ver a língua – seja de um ponto de vista diacrônico ou sincrônico – como um objeto constituído de heterogeneidade ordenada” (WEINREICH; LABOV; HERZOG, 2006, p. 35).

Essa nova maneira de encarar a língua, não mais como um sistema homogêneo e estável, mas como uma atividade social intrinsecamente variável, instável, mutante, que está permanentemente em formação e transformação, põe em evidência a seguinte constatação: “as pressões sociais estão operando continuamente sobre a língua, não de algum ponto remoto no passado, mas como força social imanente agindo no presente vivo” (LABOV, 2008, p. 21). Essa percepção de Labov parece, de certa forma, refletir uma indagação que ele, juntamente com Weinreich e Herzog, haviam feito em relação à heterogeneidade e a abordagem estrutural da língua. Para eles “se uma língua tem de ser estruturada, a fim de funcionar eficientemente, como é que as pessoas continuam a falar enquanto a língua muda, isto é, enquanto passa por períodos de menor sistematicidade?” (WEINREICH; LABOV; HERZOG, 2006, P. 35).

Além de ser uma realidade incontestável, a variação linguística, segundo Mollica (2010), “constitui fenômeno universal e pressupõe a existência de formas linguísticas alternativas denominadas variantes” (MOLLICA, 2010, p. 10). Conforme Bagno (2007), a variação “ocorre em todos os níveis da língua: variação fonético-fonológica, variação morfológica, variação sintática, variação semântica, variação lexical e variação estilístico-pragmática” (BAGNO, 2007, p. 39-40). Além disso, os fenômenos são condicionados, extralinguisticamente, o que significa que fatores sociais como origem geográfica, grau de escolarização, idade etc. têm um grande peso na ocorrência de determinadas variantes. Esse dado só faz corroborar a tese de que não dá para discutir a diversidade linguística sem levar em conta a diversidade social, até porque a língua reflete a estrutura social do falante.

Assim sendo, para a resolução de problemas relativos aos fenômenos como o do plural no sintagma nominal muitos pesquisadores têm se apoiado nos pressupostos teórico-

mulheres no uso das variantes, o comportamento de diferentes gerações no uso das variantes, o comportamento das classes sociais, dos grupos étnicos e uma série de variáveis sociais explanatórias envolvidas. Mas o grande salto que Labov dá nesse estudo de Martha's Vineyard foi justamente não se restringir a fazer unicamente essas correlações. Pelo contrário, ele faz uma interpretação do conjunto da comunidade de fala de Martha's Vineyard, sobretudo, em sua dimensão ideológica da atitude dos falantes em relação à língua. E ele observou que havia entre os moradores da ilha duas atitudes: aqueles que estavam gostando da presença dos veranistas porque iriam desenvolver economicamente a ilha e aqueles que reagiam, achando que aquela invasão estava destruindo a cultura local. Com isso, Labov estabeleceu uma correlação crucial para entender o mecanismo da mudança na ilha que foi o de que os falantes que rejeitavam a presença dos veranistas eram os que mais usavam os ditongos centralizados, como [əɪ] ao invés de /ay/. E os que eram a favor daquela tendência econômica da ilha de tornar-se um balneário, adotavam a pronúncia padrão do inglês. A importância desse estudo foi, exatamente, – e isso é importante destacar – porque Labov saiu da mera correlação mecânica do processo de variação com as variáveis sociais e fez uma interpretação de conjunto, isto é, uma interpretação mais global do contexto social em que a mudança ocorre. Isso fez com que o estudo de Martha's Vineyard se tornasse, de fato, um arquétipo para qualquer análise sociolinguística da mudança.

metodológicos da teoria da variação e mudança linguística, cujos nomes de maior relevo sem dúvida são os dos americanos Labov, Weinreich e Herzog. Ambos tiveram a percepção inaugural do quanto a língua é dotada de heterogeneidade ordenada e, por isso mesmo, não deixaram de reconhecer o quão heterogênea e ordenada é a competência dos falantes. Os estudos sociolinguísticos subsequentes a esses teóricos abraçaram a perspectiva metodológica da variação e adentraram no universo aparentemente caótico da língua falada, haurindo daí excelentes materiais que só fazem reforçar a tese de que tanto a variação no nível da língua quanto a variação motivada por algum fator de ordem social, condicionada, portanto, extralinguisticamente, apresentam ordem do ponto de vista semântico e pragmático. Além disso, são plenamente funcionais, oferecendo, conforme Bagno (2007, p. 48), “todos os recursos necessários para que os falantes interajam socialmente”.

2.1.3 Princípios teóricos fundamentais da sociolinguística

As bases sob as quais se assentam a sociolinguística são definidas de forma objetiva e direta no texto clássico *Fundamentos empíricos para uma teoria da mudança linguística*. Ali, de acordo com Lucchesi, os autores sustentam que

(I) a língua funciona enquanto muda; (II) a heterogeneidade não compromete o funcionamento da língua – um sistema homogêneo e invariável é que seria disfuncional em uma comunidade de fala culturalmente diversificada; (III) a variação faz parte do sistema linguístico, que é heterogêneo e composto por regras e unidades variáveis; (IV) a variação é potencialmente a atualização, em cada momento que se considere a língua, dos processos de mudança em curso no seu devir histórico (mudança implica variação, mas variação não implica necessariamente mudança); (V) a variação não é aleatória. A análise sincrônica dos condicionamentos estruturais e sociais da variação é capaz de revelar os mecanismos que atuam na implementação dos processos de mudança que afetam o sistema da língua; (VI) a mudança linguística pode ser estudada diretamente através da análise da variação observada em cada estado de língua (WEINREICH; LABOV; HERZOG, 2006 *apud* LUCCHESI, 2012).

Esses pressupostos definem sinteticamente o que Lucchesi (2012) chama de ruptura epistemológica oriunda, segundo ele, da contradição entre sistema e mudança. Daí, é possível estabelecer alguns princípios fundamentais do programa da sociolinguística que justificam de forma bastante eloquente as pesquisas na área, como é o caso do princípio de que a variação não é aleatória, mas inerente e, por isso mesmo, condicionada por fatores estruturais e sociais. Ora, aqueles que concebem a linguagem popular como caótica e sem regras demonstram, de

forma explícita, uma grande ignorância em relação à realidade da língua. Se a linguagem popular fosse caótica e a variação linguística fosse aleatória, o falante poderia produzir, por exemplo, no âmbito da concordância nominal no SN, tanto *os filho* quanto *filho os*. Porém, não se tem notícia, no português brasileiro, de que algum falante tenha produzido uma sentença na qual o artigo determinante venha posposto ao substantivo núcleo no sintagma nominal, como no caso do SN *filho os*. Portanto, vale reiterar que o princípio que justifica a pesquisa sociolinguística é o de que a variação na língua não é aleatória, mas ela é condicionada por fatores sociais e estruturais. E no caso do exemplo acima, o fator estrutural é que está em jogo. É a posição do constituinte dentro do sintagma nominal.

Um outro princípio é o de que a língua deve ser estudada em suas condições reais de uso e não como uma estrutura abstrata, encerrada em sua lógica interna. Até porque, a idealização é um componente que se faz presente, certamente, muito mais na pesquisa linguística que lida com a língua escrita, que é uma forma mais restrita e codificada da língua e não a realidade básica, como a oralidade. Na verdade, a escrita é uma representação da oralidade. O estudo da variação linguística pressupõe o estudo da língua em seu uso mais concreto, que é o que Labov chama de vernáculo, ou seja, a língua na sua forma de uso mais espontânea, mais direta ou, nos termos laboviano, “o estilo em que se presta o mínimo de atenção ao monitoramento da fala” (LABOV, 2008, p. 244).

Outro princípio teórico básico é o de que a língua está em processo permanente de mudança. Na verdade, é condição essencial, inerente a qualquer língua viva. Curiosamente, para aqueles que defendem que a mudança compromete o funcionamento da língua, há a constatação óbvia de que a língua, ironicamente, só para de mudar quando ela para de funcionar. Ou seja, só línguas mortas é que não mudam, como o latim, por exemplo.

Como se pode ver, o programa de pesquisa sociolinguística surge na década de sessenta como uma alternativa ao estruturalismo que era o modelo dominante até então na pesquisa linguística²¹. Para Saussure, a língua era uma estrutura encerrada em si mesma, por isso a pesquisa linguística deveria buscar entender o sistema linguístico como um todo, dispensando assim o estudo da realidade atomizada das mudanças. Na sequência do estruturalismo, sobretudo a partir dos trabalhos do círculo linguístico de praga, que foi um

²¹ Embora o estruturalismo tenha sido o programa de pesquisa hegemônico até a década de 1950, o mesmo começou, a partir dessa época, a apresentar sinais de esgotamento na medida que não conseguia responder algumas das questões que a pesquisa empírica sobre a linguagem colocava. Uma dessas questões diz respeito, exatamente, à mudança linguística. Como Saussure descartou a variação e elegeu como objeto de estudo a língua sincrônica - esse seria o objeto por excelência da análise linguística - o estruturalismo naturalmente se tornou um modelo incapaz de dar conta do fato empírico de que todas as línguas estão em permanente processo de mudança.

programa de pesquisa do estruturalismo que surgiu no final da década de 1920, vai se soar muito a ideia de funcionalidade. Então a língua não é mais um sistema, mas um sistema funcional. Um sistema que cumpre funções sociais. Jakobson (1969) fala em seis funções da linguagem. Só que na prática, o que se pode notar na análise estrutural é que a análise funcional se prende basicamente à função comunicativa da língua. Daí a ideia de um sistema unitário e homogêneo. A principal dicotomia saussuriana, aquela que opõe língua a fala, concebe a língua como um sistema abstrato e homogêneo e a fala como a manifestação concreta da língua²². No entanto, para que a língua funcione plenamente, em diferentes contextos sociais, ela tem que ser variável e heterogênea. Se a língua fosse subordinada apenas a função comunicativa, um sistema que servisse exclusivamente para a comunicação, a sua forma ideal seria aquela em que se teria uma forma para cada significado.

Um outro aspecto importante a se considerar tem a ver com a hierarquia das interações sociais. É mais que notório que a imagem social do falante é construída pela forma como ele fala. E essa é uma característica bastante visível no âmbito das relações linguísticas. Além disso, a construção da identidade social e a questão dos efeitos comunicativos, tudo isso, faz com que a variação linguística seja imprescindível para que a língua possa funcionar como ela funciona na sociedade. É pertinente destacar também que quando uma forma linguística inovadora vem de baixo para cima tende a ter uma reação por parte dos falantes da norma culta. Mas para que a mudança se implemente, ela tem que vencer essa reação.

A pesquisa sociolinguística procura também fazer um diagnóstico usando os termos variação estável e mudança em progresso ou mudança em curso. Por exemplo, se o estudo da distribuição na variação de plural no SN na estrutura da língua e na estrutura social apontar que as variantes presença e ausência vão conviver indefinidamente na língua, a concordância formal não vai desaparecer, sendo totalmente substituída pela concordância variável, então o diagnóstico é de variação estável. Se os dados empíricos apontarem na direção de que a marcação formal vai cair em desuso e vai desaparecer, então o diagnóstico seria de mudança em progresso. E aqui, vale lembrar que a possibilidade de estudar a mudança, que é um processo histórico, diacrônico, a partir de dados sincrônicos, é o que se chama, na sociolinguística, de estudo da mudança em tempo aparente. Essa perspectiva se opõe aquela visão clássica de que o estudo da mudança linguística só poderia ser feito em tempo real, comparando estágios sucessivos da língua. Portanto, ao estudar a língua do ponto de vista da

²² O estudo da língua tal qual concebida por Saussure como entidade abstrata parece ter se revigorado com Chomsky que, segundo Labov (2008, p. 218), “fortaleceu a dicotomia saussuriana ao opor a competência, ou o conhecimento abstrato das regras da língua, ao desempenho, ou seleção e execução destas regras”.

variação e da mudança, obviamente que a sociolinguística adentra na dimensão sócio-histórica do fenômeno linguístico. A língua como um fenômeno coletivo, social e histórico. Não é à toa que segundo Labov “o objeto da descrição linguística é a gramática da comunidade de fala: o sistema de comunicação usado na interação social” (LABOV, 1982, P. 18 *apud* LUCCHESI, 2011, P. 236). Por sinal, a sociolinguística define uma comunidade de fala, ou seja, uma comunidade linguística não pela semelhança entre as formas como as pessoas falam, mas por um ideal de língua comum. Para Guy (2001), por exemplo, são três os critérios que definem uma comunidade de fala, a saber: os falantes devem compartilhar traços linguísticos que sejam diferentes de outros grupos; devem ter uma frequência de comunicação maior entre si do que com outros grupos e devem partilhar normas e atitudes em relação ao uso da linguagem. Dessa forma, pela distribuição da variação na comunidade de fala é possível observar se a mudança está avançando ou se ela está em um processo de estabilização, que é um dos objetivos da pesquisa sociolinguística.

Para concluir essas considerações em torno da sociolinguística variacionista, julgamos relevante ressaltar a ideia principal de Labov, presente logo no início da obra *Padrões sociolinguísticos*, segundo a qual “não se pode entender o desenvolvimento de uma mudança linguística sem levar em conta a vida social da comunidade em que ela ocorre” (LABOV, 2008. p. 21), o que nos sugere, implicitamente, que a mudança não pode ser explicada apenas pelos chamados fatores internos, sejam eles estruturais sejam eles funcionais. Para dar conta da mudança linguística, é preciso ir ao encaixamento da variação e da mudança na estrutura social. Combinando esse condicionamento social com o condicionamento estrutural, é possível não só uma análise sincrônica da variação linguística, como também estabelecer as tendências das mudanças em curso na língua. Uma análise como a que nos propomos a fazer envolvendo o português popular e o culto, no que tange à variabilidade das marcas de plural no sintagma nominal, não pode prescindir de uma interpretação que leve em conta a ação simultânea de ambos os condicionamentos na ocorrência do fenômeno, sob pena de se restringir única e exclusivamente a algumas generalizações um tanto precipitadas de que haveria, por exemplo, uma tendência para a passagem gradual de um sistema com concordância para um sistema sem concordância nominal, notadamente entre os falantes da norma popular, ou que as mulheres fazem mais concordância nominal formal do que os homens. Muitas pesquisas têm mostrado o contrário, evidenciando assim que só é possível esse tipo de inferência quando se analisa a inserção do falante em cada realidade histórica e social concreta em que a mudança esteja acontecendo.

Na próxima seção, é possível visualizar melhor a aplicabilidade da teoria variacionista aos estudos da concordância nominal de número no português brasileiro.

2.2 ESTUDOS SOBRE A VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO PORTUGUÊS DO BRASIL

O português brasileiro vem sendo objeto de estudo já há algum tempo. Dialetólogos, filólogos e linguistas têm se debruçado sobre os fenômenos linguísticos que pululam nas mais diversas comunidades de fala espalhadas por esse enorme país de dimensões continentais chamado Brasil. Na obra *Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica*, organizada por Sebastião Votre e Cláudia Roncarati (2008, p. 16), consta uma entrevista com o professor Naro na qual ele considera que o português que aqui chegou, no início da colonização, já trazia as marcas da variação.

Além disso, em tal entrevista, Naro afirma que a população portuguesa que veio para o país, nessa época, era, indiscutivelmente, na sua grande maioria, analfabeta, falante do português popular de Portugal que era, por sua vez, uma variedade, certamente, diferente da norma culta do português europeu daquela época. O que sugere que o contato linguístico inicial dessa variedade com as línguas indígenas e depois com as africanas tenha sido regido pelas marcas da variabilidade. Obviamente, que esses contatos, incluindo, posteriormente, aqueles que ocorreram com outras variedades oriundas de línguas europeias e asiáticas – restrito a determinadas regiões do país que receberam levadas de imigrantes internacionais a partir do final do século XIX – deram ao português brasileiro uma fisionomia peculiar que o distingue de longe da matriz europeia. Basta, para tanto, comparar alguns fenômenos de natureza fonética, semântica ou morfossintática para se constatar logo o domínio geográfico de cada uma das variedades: a portuguesa e a brasileira.

Entre os vários fenômenos linguísticos caracterizadores de cada um dos macros espaços em que a língua portuguesa viceja soberana, um deles tem recebido um tratamento especial nas últimas décadas na pesquisa sociolinguística brasileira. Trata-se da concordância gramatical de número plural entre os elementos flexionáveis do sintagma nominal.

O pioneirismo nesse estudo se deve a Maria Marta Scherre. O trabalho inaugural da autora começa a ganhar notoriedade em 1976 numa pesquisa sobre a concordância de número no SN na área urbana do Rio de Janeiro. Trabalho esse realizado junto com Maria Luiza Braga. Em 1978, Scherre pesquisa a regra de concordância de número no sintagma nominal em português e a partir daí não para mais de produzir estudos cada vez mais significativos.

Mas foi a tese sobre a *reanálise da concordância nominal em português*, de 1988, que mudou radicalmente o rumo das discussões acerca da influência das variáveis linguísticas e extralinguísticas sobre a marcação de plural nos constituintes do SN. Um estudo extremamente minucioso, repleto de dados quantitativos que além de corroborar o estado de variabilidade em que se encontra o fenômeno da concordância nominal no português brasileiro, aponta para uma ação pujante e simultânea de algumas variáveis que ora favorecem, ora desfavorecem a variação.

Scherre não foi a única a adentrar o universo variacionista do sistema flexional nominal do português brasileiro. As pesquisas envolvendo a concordância entre os elementos do sintagma nominal cobrem um número bastante significativo de dados analisados em diferentes regiões do país. Num intervalo aproximado de quatro décadas, é possível contabilizar uma gama muito expressiva de dissertações de mestrado e teses de doutorado, sem contar os inúmeros artigos científicos publicados em revistas especializadas, todos empreendidos à luz dos postulados da sociolinguística laboviana.

Para melhor visualizar essa fecunda atividade produtiva de descrição, análise e sistematização do fenômeno em questão, parece-nos relevante destacar alguns trabalhos realizados no Brasil, nos últimos anos, sem deixar de mencionar, a seguir, aqueles que marcaram o início dos estudos sociolinguísticos sobre a concordância nominal.

Destarte, além das pesquisas de Scherre supracitadas, e que serão oportunamente retomadas para embasar as discussões de natureza linguísticas e extralinguísticas do presente estudo, são dignos de nota, igualmente, a dissertação de Braga (1977), na qual ela analisou a fala de sete moradores do triângulo mineiro, de classe média e baixa. A dissertação de Ponte (1979) que apresenta análise da fala de vinte moradores de Porto Alegre, todos analfabetos. O estudo de Terezinha Nina (1980), que em sua dissertação de mestrado trabalhou dados de vinte informantes analfabetos de Bragantina (PA). A pesquisa de Guy (1981) que, por sua vez, analisou dados da pesquisa Competências Básicas, envolvendo vinte informantes cariocas, semi-escolarizados, da área urbana. A análise de Dias (1993), que envolveu dados de vinte falantes de Brasília, com quatro anos de escolarização, sendo dez da área rural e dez da área urbana. Além da pesquisa de Campos e Rodrigues (1993) que analisaram dados de um corpus do projeto NURC, envolvendo informantes universitários de cinco capitais brasileiras²³.

²³ O levantamento desses estudos foi feito com base em Scherre (1994, p. 2), no artigo intitulado *Aspectos da concordância de número no português do Brasil*, publicado na Revista Internacional de Língua Portuguesa (RILP).

Somadas a essas investigações preliminares acerca da variação na concordância nominal, outras pesquisas surgiram no cenário acadêmico, dando maior visibilidade à ocorrência de tal fenômeno em outras partes do país. São os casos, por exemplo, dos trabalhos dos seguintes autores:

- ✓ Marisa Fernandes (1996) que, em sua dissertação de mestrado, analisou a fala de moradores da Região Sul do Brasil, tomando por base as variáveis sociais: sexo, escolaridade, etnia e nível de formalidade;
- ✓ Raimunda Carvalho (1997), cuja pesquisa centrou na investigação da fala de informantes de Rio Branco (AC), de classe social baixa, considerando o sexo e a escolaridade;
- ✓ Hebe Carvalho (1997), que analisou a fala de moradores de João Pessoa (PB), correlacionando o fenômeno da variação na concordância nominal aos fatores idade, escolaridade, sexo/gênero;
- ✓ Norma Lopes (2001), que em sua tese de doutorado analisou as falas popular e culta de moradores de Salvador (BA), tomando por base a seguinte configuração social: quatro diferentes faixas etárias, três graus de escolarização, dois gêneros e dois diferentes grupos de etnia;
- ✓ Elza da Silva Bueno (2002), que analisou a concordância gramatical de número entre 144 informantes dos sexos masculino e feminino que trabalham no corte de cana de açúcar, de seis localidades diferentes do estado de São Paulo. Esses falantes têm pouquíssima ou nenhuma escolaridade e foram distribuídos em três faixas etárias;
- ✓ Odette de Souza Campos e Ângela Rodrigues (2002), que pesquisando dados coletados no projeto NURC (Norma Urbana Culta), analisaram a fala de informantes com nível superior completo ou em andamento, envolvendo fatores como idade, origem geográfica e nível de escolaridade;
- ✓ Patrícia Ribeiro de Andrade (2003), que analisou também a recorrência de variantes na concordância nominal, no dialeto de Helvécia (BA), com base em um *corpus* constituído de dezoito inquéritos, com informantes de três faixas etárias de ambos os sexos;
- ✓ Lília Soares Miranda Santos (2010) que, a propósito de dissertação de mestrado, analisou a fala de informantes de Pedro Leopoldo (MG), levando em consideração fatores como idade, escolaridade, sexo e classe social;
- ✓ Janaína Biancardi da Silva (2011), que fez a análise da fala de informantes de Vitória (ES), observando os fatores idade, escolaridade e sexo;

- ✓ Sílvia Figueiredo Brandão (2011), cuja pesquisa envolveu a fala de informantes de duas variedades urbanas do português: uma de Niterói-RJ, Brasil, e outra de São Tomé e Príncipe. Na análise, o autor considerou fatores como sexo, escolaridade e idade;
- ✓ Flávio Marcelo Bueno de Castro e Vinícius Carvalho Pereira (2012), que se debruçaram na análise da fala de informantes com nível superior completo da cidade de Cuiabá-MT;
- ✓ Simone Daise Schneider (2012), que apresenta uma análise muito interessante da concordância nominal na fala de crianças de 3 a 6 anos de idade do município de Novo Hamburgo (RS) e que frequentam a educação infantil. Os informantes são oriundos de famílias de escolaridade média e superior e de classe social média-alta;
- ✓ Vívian Antonino (2012) que, em sua tese de doutorado, investigou a concordância nominal em predicativos do sujeito e em estruturas passivas do português popular de Salvador;
- ✓ Flávia Santos Martins (2013), que analisou a fala de habitantes do alto Solimões (AM), apoiando-se nas seguintes variáveis extralinguísticas independentes: idade, escolaridade, sexo, diatopia, ocupação, mobilidade e localismo.
- ✓ Maria Aparecida Souza Guimarães (2014), que analisou a fala de conquistenses cujo vernáculo a pesquisadora considerou como Português Popular, tendo como parâmetro: idade, escolaridade, sexo.

Todos esses estudos são unânimes em apontar variação do plural no SN, sobretudo entre as classes sociais menos privilegiadas, semi-escolarizadas ou sem nenhuma escolarização. Além do mais, em relação à realidade sociolinguística brasileira específica, as pesquisas em diferentes regiões do país evidenciam que:

O fenômeno da variação na concordância de número no português falado do Brasil, longe de ser restrito a uma região ou classe social específica, é característico de toda comunidade de fala brasileira, apresentando diferenças mais de grau do que de princípio, ou seja, as diferenças são mais relativas à quantidade de marcas de plural e não aos contextos linguísticos nos quais a variação ocorre (SCHERRE, 1994, p. 2).

A autora ainda destaca que o fenômeno da variabilidade de plural no sintagma nominal no português brasileiro “pode ser caracterizado como um caso de variação linguística inerente, tendo em vista que ocorre em contextos linguísticos e sociais semelhantes e apresenta tendências sistemáticas de variação altamente previsíveis” (SCHERRE, 1994, p. 2).

Dos inúmeros estudos sobre a variação na concordância nominal no português, elencados acima, procuramos nosiliar àqueles que apresentam contribuições importantes

para a nossa pesquisa cujo foco de análise envolve a comparação entre as normas popular e culta da comunidade de fala de Vitória da Conquista, no que diz respeito à variação na marcação de plural no sintagma nominal²⁴.

Não dá para discutir a variação na indicação de pluralidade no sintagma nominal sem nos reportar a tese de Scherre *Reanálise da concordância nominal em português*. Um legítimo tratado sociolinguístico quantitativo, essa tese é uma obra que veio preencher uma enorme lacuna deixada nos estudos linguísticos relativos à variabilidade de marcas de plural no SN. O título, por si só, justifica tal asserção, mas nunca é demais lembrar as pesquisas nessa área que antecederam o trabalho da autora e que revelaram a necessidade de uma análise mais conclusiva do fenômeno.

Assim, vale destacar que a pesquisa feita por Braga e Scherre em 1976 é tida como um dos primeiros estudos realizados sobre a concordância entre os elementos do sintagma nominal. Elas analisaram dados de sete falantes residentes no Rio de Janeiro, porém de classe social e origem geográfica diferentes. Tal como nos trabalhos de Braga (1977), de Scherre (1978), Guy (1981), que consideraram apenas a posição linear, bem como os estudos que consideraram apenas a classe gramatical – o que inclui Cedergren (1973) com o estudo do espanhol do Panamá –, assim como aqueles estudos que consideraram ambos os fatores como variáveis separadas, como é o caso do trabalho de Poplack (1981) com o espanhol de Porto Rico ou os que consideraram haver uma equivalência entre essas variáveis, por exemplo, Guy (1981); o que se observou em todos esses estudos em relação a tais variáveis linguísticas independentes é que as abordagens em torno delas se mostraram insuficientes para uma interpretação mais adequada do fenômeno da variação na concordância nominal do português brasileiro.

Nesse sentido, a originalidade do trabalho de Scherre em se propor a reanalisar de forma minuciosa a concordância em todos os níveis de análise estrutural e social possíveis, apresentando, no caso específico das variáveis posição e classe gramatical, uma nova

²⁴ O fenômeno linguístico por nós estudado, e que servirá de base para o estudo contrastivo entre as normas popular e culta, tem recebido ao longo dessa dissertação algumas denominações que podem, a primeira vista, exprimir uma certa indefinição quanto ao real objeto de análise. Assim, é possível encontrar, por exemplo: *a variação de plural na concordância nominal*, *a variação na marcação de plural no sintagma nominal*, *a concordânciagramatical de número plural entre os elementos flexionáveis do sintagma nominal*, *a variabilidade de marcas de plural no SN* etc. que, na verdade, só fazem reforçar mais ainda a ideia de variação nos constituintes que compõem o sintagma nominal. Scherre (1988, p. 62) chega a questionar o uso do termo concordância para se referir à variação, considerando-o inadequado, tendo em vista aqueles casos de variabilidade em que apenas um elemento recebe marca formal de plural, como no exemplo: *as perna toda marcada*, indo, portanto, segundo ela, de encontro ao princípio de que a “concordância implica harmonia formal em pelo menos dois elementos de uma dada construção”. Mas, mesmo assim, a própria autora reconhece que não tem como não referir a casos como o desse exemplo a presença de concordância.

abordagem que as consideram não mais isoladamente, mas na interrelação entre ambas e na relação que se estabelece entre os determinantes e o núcleo do sintagma nominal. Ao estabelecer, portanto, esse cruzamento entre as variáveis em questão como critério fundamental para uma análise mais profunda, a autora evidencia com isso a incompletude das análises que antecederam a sua tese, já que as mesmas acabavam encobrindo regularidades linguísticas importantes.

A pesquisa de Scherre envolveu um corpus constituído de 64 gravações de 60 minutos cada envolvendo 64 informantes, sendo 48 adultos (15-71 anos) e 16 crianças (7-14 anos), radicados no município do Rio de Janeiro. Gravações essas extraídas do banco de dados do *Corpus Censo* vinculado à faculdade de Letras da UFRJ.

A amostra dos falantes envolvidos na pesquisa foi estratificada pela autora em função de três características sociais, a saber: sexo, anos de escolarização (1 a 4, 5 a 8 e 9 a 11) e faixa etária (7-14, 15-25, 26-49 e 50-71). Os dados foram caracterizados por Scherre considerando como objeto de estudo todos os sintagmas nominais “que apresentasse pelo menos uma marca formal ou semântica de plural que implicasse necessariamente outra ou outras marcas formais de plural nos elementos flexionáveis do SN” (SCHERRE, 1988, p. 31-32). Sendo assim, todos os sintagmas nominais cuja variação subvertesse a prescrição normativa da gramática tradicional se constituíram em tal objeto. Com isso, foi possível determinar de forma bastante criteriosa que a marca explícita de plural poderia ser encontrada em todos os elementos flexionáveis do SN, em alguns elementos flexionáveis do SN, em apenas um dos elementos flexionáveis do SN e, até mesmo, em nenhum dos elementos flexionáveis do SN desde que ocorresse um sintagma contendo um numeral como primeiro elemento seguido de outros elementos não marcados. É óbvio que o rigor científico exige uma verdadeira “depuração” dos dados a fim de se obter resultados cada vez mais confiáveis e conclusivos. E Scherre (1988) não fugiu em momento algum dessa lógica, sobretudo, ao considerar como irrelevantes estruturas sintagmáticas que não se encaixavam no princípio norteador de qualquer pesquisa que lida com o fenômeno da variação linguística: a frequência em que se manifesta tal fenômeno. Assim, é perfeitamente justificável, na análise da autora, a exclusão dos sintagmas que fugiam da posição canônica, como em *tios, eu tenho dois*. Os casos em que a autora considera cristalizados, como *os trapalhões*, assim como aqueles em que apresentam núcleo nominal invariável: *costas, férias* etc. Além disso, foram excluídos também casos em que os elementos sintagmáticos ocorriam em situação de neutralização, como em: *as meninassão, as botascheias, osjogadores*.

O levantamento quantitativo dos dados passíveis de análise, conforme projetado por Scherre (1988), redundou num total de 7.193 sintagmas nominais, sendo que 6.027 produzidos pelos informantes adultos e 1.166 pelas crianças. De posse desses dados, ela estabeleceu alguns objetivos com a análise do *corpus*, sendo que um deles é digno de nota, pois sintetiza de forma bastante eloquente o grau de profundidade com que a autora trata um fenômeno que para os não iniciados teria uma explicação aparentemente óbvia: a retenção de plural apenas na primeira posição do sintagma por conta do caráter redundante da concordância nominal de número no português. Scherre (1988) vai mais longe, perpassa essa conclusão ingênua, assegurando:

[...] que existe um sistema gerenciando a variação, sendo, portanto, possível estabelecer-se com relativa precisão em que circunstâncias linguísticas haveria a ocorrência de uma estrutura sintagmática com todas as marcas de plural ou com apenas algumas delas. Consequentemente, é possível prevê em que estruturas linguísticas os falantes são propensos a colocar ou não todas as marcas formais de plural” (SCHERRE, 1988, p. 15).

A autora trabalhou com a concordância nominal sob duas perspectivas: atomística e não atomística. Na primeira, cada elemento flexionável do sintagma se constitui em unidade de análise. Na segunda, ao contrário, a unidade analisada é a própria estrutura sintagmática. As variantes consideradas em tais perspectivas foram presença versus ausência de marcas formais de plural.

A análise quantitativa dos dados no contexto atomístico envolveu um total de 13.229 constituintes sintagmáticos nominais extraídos de 7.193 sintagmas. Já na perspectiva não atomística, foram analisados 948 sintagmas apenas dos informantes adultos. Para a obtenção de resultados probabilísticos mais precisos, Scherre (1988) extrapolou o número de variáveis independentes, elegendo onze linguísticas e seis sociais na análise atomística e nove linguísticas e três sociais na análise não atomística.

Vale destacar que, ao trabalhar em duas frentes de análise: atomística e não atomística, elegendo as variáveis potencialmente responsáveis pelo mecanismo de flexão nominal, a autora só faz ratificar mais ainda a inegável conclusão de que a concordância nominal no SN é, de fato, uma regra variável, sendo, portanto, determinada pela atuação de diversas variáveis, entre as quais se encontram aquelas que na reanálise de Scherre exerceram uma forte influência sobre a concordância. Ela, inclusive, apresenta as seguintes que serviram de base para sua análise na perspectiva atomística: saliência fônica; dimensões processos e tonicidade; relação entre os elementos não nucleares em função do núcleo e posição dos

elementos nucleares no SN; marcas precedentes em função da posição; contexto fonético/fonológico seguinte; função resumitiva do SN; formalidade dos substantivos e adjetivos; grau dos substantivos e adjetivos; animacidade dos substantivos. Na abordagem não atomística, a autora trabalha com as seguintes variáveis: pluralidade do contexto; configuração sintagmática do SN; grau/formalidade do SN; pluralidade do SN; função textual; localização do SN; saliência fônica: dimensão processos.

Embora pareça um tanto exaustivo discorrer sobre cada uma dessas variáveis, vamos expor aqui ao menos algumas considerações acerca daquelas que se mostraram bastante atuantes no processo de condicionamento do fenômeno em questão. E nesse sentido, Scherre foi bastante minuciosa na descrição e explicação de cada uma das variáveis, o que aumenta, consideravelmente, a responsabilidade em não omitir aspectos relevantes apresentados pela autora em sua tese.

Assim, entre as variáveis que serviram de base para a análise da concordância nominal na perspectiva atomística, merece destaque especial na obra de Scherre o princípio da saliência fônica. De acordo com a autora, esse princípio “consiste em estabelecer que as formas mais salientes, e por isso mais perceptíveis, são mais prováveis de serem marcadas do que as menos salientes” (SCHERRE, 1988, p. 64). Introduzido de forma inaugural nos estudos de Naro e Lemle, em meados da década de 70, para explicar aspectos do funcionamento sincrônico do português do Brasil, sobretudo no que se refere à concordância verbal, essa variável foi bastante explorada pela autora nos seus trabalhos iniciais sobre a concordância de número entre os elementos do SN. Na sua primeira versão, a saliência era considerada apenas na dimensão que envolvia a quantidade de diferenciação de material fônico na relação singular/plural. Em seguida, Naro propôs que mais outra, a tonicidade dos itens lexicais singulares, fosse estabelecida no eixo da saliência fônica. Em sua reanálise, Scherre propôs, igualmente, uma terceira, que seria o número de sílabas. Portanto, em sua tese, a autora considerou tal princípio sob as três dimensões estabelecidas: processos morfofonológicos de formação de plural (também conhecido como diferenciação de material fônico), tonicidade da sílaba dos itens lexicais singulares e número de sílabas dos itens lexicais singulares. Para cada uma dessas dimensões, foram estabelecidas algumas subdivisões, obedecendo assim a uma escala de níveis hierárquica na qual foi possível obter resultados mais conclusivos acerca daqueles fatores que mais favorecem a presença da concordância em função do grau de saliência fônica. É certo que os resultados obtidos por Scherre em sua pesquisa revelaram que o número de sílabas não exerceu influência sobre a concordância de número entre os elementos do SN, o que levou a autora a reconsiderar a

saliência fônica sob duas dimensões: processos e tonicidade. Porém, não se dando por satisfeita com as análises em que ambas as dimensões eram consideradas como variáveis separadas, Scherre se propôs a fazer uma análise alternativa na qual processos e tonicidade se fundiram em uma só variável, contendo o seguinte grupo de fatores:

- 1) plural duplo (oxítono ou paroxítono);
- 2) itens terminados em -l (oxítono ou paroxítono);
- 3) itens terminados em -R (oxítono ou paroxítono);
- 4) itens terminados em -ão (oxítono);
- 5) itens terminados em -S (oxítono);
- 6) regular oxítono;
- 7) regular paroxítono e
- 8) regular proparoxítono.

Os resultados obtidos com a nova variável confirmou, no estudo de Scherre, o princípio de que as formas mais salientes, portanto mais perceptíveis, são mais marcadas conforme os pesos relativos a seguir: itens de plural duplo 0,80; itens terminados em (-l, -R, -ão, -S) variaram entre 0,56 e 0,69; itens regulares oxítonos e monossílabos tônicos 0,38; itens regulares paroxítonos 0,21 e regulares proparoxítono 0,17.

Outra variável que se mostrou bastante relevante no processo de análise tem a ver com *marcas precedentes em função da posição*. Na verdade, tanto essa variável quanto a que se refere à *relação entre os elementos não nucleares em função do núcleo e posição dos elementos nucleares no SN* são resultados de uma hábil combinação da autora ao constatar que as variáveis: Classe gramatical, Posição e Marcas precedentes não davam conta de explicar o fenômeno da variação da concordância nominal em sua totalidade. Daí a transformação dessas três variáveis em duas por meio do cruzamento entre elas. A variável *marcas precedentes e posição*, por exemplo, permitiu verificar que a ocorrência de marcas de plural em um item anterior leva a marcas no item seguinte e a ausência de marcas em um item anterior leva à ausência no item seguinte. Ou seja, marcas levam a marcas e zeros levam a zeros. A autora chega a afirmar que essa mesma lógica é aplicada à variável *pluralidade do contexto*, havendo, portanto uma explicação respaldada no “Princípio do Processamento com Paralelismo, que subjaz à tendência de formas semelhantes se agruparem, através de processos mentais associativos presentes no desempenho linguístico” (SCHERRE, 1988, p. 511).

No que concerne à variável *relação entre os elementos não nucleares em função do núcleo e posição dos elementos nucleares no SN*, Scherre (1988, p. 219) observou em seu estudo que os constituintes sintagmáticos não nucleares antepostos ao núcleo recebem mais indicação de pluralidade (0,87) do que os pospostos (0,25), o que evidencia maior grau de

coesão entre os elementos que antecedem a base nuclear do sintagma nominal. Além do mais, essa variável aponta para a necessidade de se repensar os estudos em que a posição linear tem sido considerada fundamental para se compreender o fenômeno da variação de marcas de plural entre os elementos do SN. Até porque, a primeira posição da estrutura sintagmática sempre foi vista como o fator mais importante para a marcação de plural, justamente pelo fato da variável posição ser considerada em função do local que o elemento sob análise ocupa no SN. A proposta da autora rompe com essa perspectiva de análise, demonstrando que “o elemento nominal não nuclear pode estar até na terceira posição, mas, se estiver anteposto ao núcleo, apresenta também mais chances de ser marcado do que o posposto” (SCHERRE, 1988, p. 220). Ademais, verificou-se que o substantivo da primeira posição chega a ser menos marcado do que outras classes antepostas ao núcleo que ocupam primeira e segunda posição no SN.

Em relação à variável *contexto fonético/fonológico seguinte*, Scherre lembra que a variação na concordância de número no SN se constitui, antes de tudo, em um fenômeno de natureza morfossintática, portanto, o seu “funcionamento é explicado por forças alheias ao sistema linguístico” (SCHERRE, 1988, p. 515). Daí, a quase irrelevância de tal variável cuja influência se restringe, ligeiramente, aos traços consonantais [surdo], [oral] e [velar].

As variáveis *formalidade dos substantivos e adjetivos* e *grau dos substantivos e adjetivos* atuam, basicamente, em situações discursivas de pouca formalidade²⁵. E a autora mostra com grande propriedade que a presença tanto de gírias quanto de um item no diminutivo, na estrutura sintagmática, desfavorece consideravelmente a concordância nominal.

A variável *animacidade dos substantivos* foi também explorada na reanálise de Scherre (1988), tomando por base o princípio da saliência fônica. Assim, os elementos do SN com traço [+humano] se apresentam, no contexto discursivo, mais salientes, isto é, mais perceptíveis, favorecendo, portanto, a presença de marcas de plural (0,55). Por outro lado, os elementos com o traço [-humano] desfavorecem a marcação (0,45).

Para explicar a atuação de algumas variáveis no fenômeno da variação da concordância nominal, tais como *função resumitiva e localização do SN*, Scherre se respaldou

²⁵ Scherre (1988, p. 268) apresenta como itens com pouca formalidade ou, mais precisamente, informais alguns substantivos como *cara, outro* (de emprego substantivado) e aqueles usados como gírias, tais como: *besteira, bobagem, grilo, lance, papo, merda, goró*. Foram inclusos, igualmente, no rol da informalidade lexical os diversos usos da palavra *coisa*, quando desprovida de qualquer função sintática, bem como, estruturas do tipo *um montão de, uma pá de, um bando de, um monte de*. Os pesos probabilísticos acima referem-se aos dados dos adultos. Em relação às crianças, houve uma pequena variação – 0,46 [-humano] e 0,54 [+humano] – em termos proporcionais aos adultos.

no princípio da baixa tensão ou princípio da topicidade. De acordo com tal princípio, há, no discurso, uma forte tendência para os elementos não centrais aparecerem menos marcados. Os sintagmas resumitivos ou de função fática, por exemplo, se encaixam perfeitamente nessa lógica uma vez que eles tendem a ocorrer à direita da oração, uma posição considerada de baixa polaridade nos termos de Naro e Votre, como lembra a autora.

A variável *configuração sintagmática*, numa perspectiva não atomística, foi retomada por Scherre (1994, p. 7) que chegou à conclusão de que as estruturas sintagmáticas que apresentam [+artigo definido na primeira posição, + substantivo ou categoria substantiva da na última posição] “são duas forças poderosas no sentido de provocarem a presença de todas as marcas de plural em todos os elementos flexionáveis do SN”.

É importante frisar também que a relação entre variação e mudança entrou na pauta das discussões em torno do objeto a que a pesquisadora se propôs a reanalisar e a reinterpretar minuciosamente. É no contexto desse debate que emergem as reflexões relacionadas às variáveis sociais: *sexo*, *grau de escolarização* e *faixa etária*; bem como, o comportamento dessas variáveis na variação linguística estável e na mudança em progresso, consideradas pela autora como significativas para a interpretação do fenômeno pesquisado. Scherre (1988) conclui que a concordância de número entre os elementos do sintagma nominal é um fenômeno linguístico variável que apresenta duas tendências inequívocas em dois subagrupamentos distintos de falantes:

- ✓ variação estável com gradação etária nos falantes de ambiente não humilde e concordância alta;
- ✓ processo de mudança linguística em direção a um sistema sem marcas nos falantes de ambiente humilde e concordância baixa. (SCHERRE, 1988, p. 482)

Vale lembrar que, no fenômeno estudado pela autora, as variáveis sexo e grau de escolarização desempenharam importantes funções, permitindo, por exemplo, assegurar com relativa precisão que as mulheres e as pessoas mais expostas à escola tendem a se aproximar mais das formas de prestígio. Já a variável faixa etária parece ter sido aquela que exerceu menos influência, considerando o conjunto dos dados analisados pela autora.

A pesquisadora analisou também três variáveis que ela considera não convencionais, com as seguintes denominações: mercado ocupacional, mídia e sensibilidade linguística. A primeira tem a ver com a relação entre a atividade profissional desenvolvida pelo falante e o uso das formas linguísticas de prestígio que se manifesta na fala como resultado da necessidade imposta pela ocupação. A segunda se relaciona com o grau de consciência do

falante no contexto da diversidade linguística. Quanto a terceira variável, o que está em jogo é o grau de exposição do falante aos meios de comunicação de massa. Com o resultado da análise, Scherre (1988) chega à conclusão de que a variável de maior influência na ocorrência de marcas formais de plural é o *mercado ocupacional*, com peso relativo de 0,69. A variável *sensibilidade linguística* também apresentou resultados relativamente significativos, de 0,57, já que, segundo a autora, “as pessoas com maior sensibilidade fazem mais concordância” (SCHERRE, 1988, p. 492). A variável *mídia* foi a que apresentou uma influência quase que irrelevante, com peso de 0,53.

Ampliando o debate em torno dos estudos realizados sobre a variação na concordância nominal, notadamente as variáveis independentes envolvidas nas análises, optamos por dar continuidade a essas discussões reportando-nos às pesquisas ulteriores à tese de Scherre. Sendo assim, parece-nos pertinente passar à descrição da pesquisa de Marisa Fernandes (1996), destacando os principais resultados encontrados por ela na análise dos dados, em sua dissertação de mestrado. O interesse pelo trabalho da autora se justifica uma vez que ela elegeu duas situações de uso da concordância nominal de número: uma formal e outra informal, o que pressupõe dados linguísticos tanto da norma popular do português brasileiro quanto da norma culta. E essa relação comparativa ou contrastiva entre ambas as variedades muito nos interessa, tendo em vista o propósito do nosso estudo.

Em relação à situação informal, a autora utilizou o *corpus* do projeto VARSUL no qual foi possível analisar a fala de 47 informantes da região sul do Brasil a partir da coleta de 5.424 SNs. Os falantes foram estratificados, obedecendo a seguinte distribuição: 12 moradores de Florianópolis-SC (6 homens e 6 mulheres de etnia açoriana); 12 de Chapecó-SC, (6 homens e 6 mulheres de etnia italiana); 12 de Panambi-RS (6 homens e 6 mulheres de etnia alemã) e 11 informantes de Irati-PR (5 homens e 6 mulheres de etnia eslava). Nessa análise, a autora considerou três níveis de escolaridade: primário, ginásio e colegial e duas faixas etárias: 25-49, mais de 50 anos.

No que tange à situação formal, Fernandes (1996) analisou a fala de 19 informantes de algumas regiões brasileiras, tomando por base comentários esportivos (368 dados de 5 comentaristas esportivos de programas televisivos), entrevistas em televisão (415 SNs de 4 entrevistadores de canais de televisão variados) e defesas de dissertações de mestrado (738 dados de SNs envolvendo 10 defesas de diferentes áreas).

As variáveis linguísticas consideradas pela autora foram: distribuição dos elementos do SN em função de sua posição e de sua relação com o núcleo, posição linear do elemento no SN, marcas precedentes, saliência fônica (nas dimensões: processos morfofonológicos de

formação de plural e tonicidade dos itens lexicais), contexto fonético-fonológico seguinte, classe gramatical e grau e animacidade dos substantivos e adjetivos. Em relação às variáveis extralinguísticas, a autora considerou nível de escolarização, sexo, idade, etnia e formalidade da situação de fala.

De modo geral, ela chegou à conclusão de que a posição que o constituinte sintagmático ocupa em relação ao núcleo influencia a presença ou a ausência de marcas formais de plural no sintagma nominal. O que significa dizer que os elementos não nucleares antepostos ao núcleo são muito mais marcados do que os pospostos. A autora apresenta, inclusive, os seguintes pesos relativos: 0,85 para elementos não nucleares na primeira posição, 0,72 para os da segunda posição; 0,32 e 0,25 para os pospostos ao núcleo na segunda e demais posições, respectivamente. A classe gramatical dos constituintes do SN não se mostrou muito relevante no que concerne à indicação de pluralidade. Quanto a marcas precedentes, a autora ao correlacioná-las à variável posição, evidenciou que a presença de marcas formais de plural na primeira posição leva à marcação nos elementos seguintes (0,61) e zero formal desfavorece a concordância plural (0,10). Em outras palavras: marcas levam a marcas e zero leva a zero.

A saliência fônica foi outra variável explorada por Fernandes (1996) e que, segundo a autora, as formas mais perceptíveis, por isso mais salientes, são mais marcadas. Os nomes, por exemplo, que permitem a formação de plural duplo e aqueles que terminam em -L tendem a favorecer muito mais a marcação de plural, conforme os pesos relativos apresentados pela autora: 0,81 (plural duplo) e 0,77 (itens terminados em -L).

Enquanto a animacidade dos substantivos e adjetivos não teve relevância nenhuma nos dados analisados pela autora, o grau, por sua vez, se apresentou relevante na sua forma normal, com peso de 0,51, apresentando, portanto, mais marcas formais de plural do que os constituintes nos graus diminutivo e aumentativo (0,22).

Na análise do contexto fonético-fonológico seguinte, o que se observou foi uma leve influência da pausa como fator condicionante da marcação de plural (0,56) em relação à consoante (0,48) e à vogal (0,51).

No contexto das variáveis sociais sob análise, Fernandes (1996) confirmou o que as pesquisas sociolinguísticas já haviam constatado, ou seja, quanto maior o nível de escolaridade dos falantes maior é a incidência de marcação de plural na estrutura sintagmática. O resultado obtido por ela, na situação de informalidade, ilustra com perfeição tal conclusão: colegial (0,66), ginásio (0,48) e primário (0,32).

Das demais variáveis sociais analisadas pela autora, é possível apresentar, ainda que de forma resumida, a seguinte conclusão a que ela chegou: a maior aplicação da regra de

concordância, em relação ao sexo, se deu entre as mulheres, com o peso correspondente a 0,53 contra 0,46 dos homens; quanto à idade, os mais velhos – acima de 50 – fizeram mais uso de marcas formais de plural (0,55), superando os mais jovens (0,45); para a variável etnia, a autora concluiu que os descendentes de origem eslava e alemã aplicaram mais a regra de plural, com peso relativo de 0,55 para ambos, do que os de origem italiana e açoriana. O que se pode inferir dos estudos de Fernandes, além dos resultados aqui apresentados, é que a aplicação da regra de concordância nominal está vinculada a uma situação muito mais formal. O que pode ser confirmado com os seguintes pesos por ela encontrados: 0,82 para a situação mais formal e 0,43 para a situação informal.

A pesquisa realizada por Hebe Carvalho (1997) se mostrou por demais esclarecedora no que diz respeito a variável dependente aqui discutida. A investigação feita por ela envolveu sessenta informantes de João Pessoa (PB), tomando por base entrevistas feitas para o Projeto VALPB. A estratificação da amostra foi distribuída em três faixas etárias: 15 a 25 anos, 26 a 49 anos e acima de 50 anos; nos níveis de escolaridade seguintes: analfabetos, 1 a 4 anos, 5 a 8 anos, 9 a 11 anos e nível universitário e, finalmente, conforme o sexo: homem e mulher.

Dentre as inúmeras variáveis linguísticas analisadas pela autora, merecem destaque, enquanto condicionadores do fenômeno da variação de plural no SN, as seguintes: classe e posição em relação ao núcleo, marcas precedentes em função da posição e saliência fônica. O resultado obtido pela autora para a primeira variável reforça a constatação de que os elementos à esquerda do núcleo recebem muito mais marcas formais de plural do que os da direita.

Vale lembrar que a autora trabalhou com duas variáveis extralinguísticas: escolaridade e faixa etária, evidenciando o peso que tem um maior nível de escolarização no uso das formas de prestígio, como é o caso do plural nos constituintes flexionáveis do SN. Em se tratando da faixa etária, foi possível observar, na análise da autora, que as maiores ocorrências da regra de plural se deram entre os informantes de 15 a 25 (0,61) e os acima de 50 (0,48).

O estudo de Raimunda Carvalho (1997) sobre a variação na concordância nominal, envolvendo 24 falantes da cidade de Rio Branco (AC) - de ambos os sexos e de três níveis de escolaridade: analfabetos, 1ª à 4ª série, 5ª à 8ª série e com idades entre 20 a 35 anos – evidenciou, a partir dos resultados das análises linguísticas e extralinguísticas, uma tendência maior por parte dos informantes para o uso das marcas formais de plural, chegando a 67% dos casos. Para se chegar a tal conclusão, a autora considerou mais detidamente a influência de três fatores linguísticos: saliência fônica, posição e classe gramatical, assumindo, nitidamente,

a mesma posição de Scherre (1988), Fernandes (1996) e Hebe Carvalho (1997) de que formas mais salientes tendem a favorecer a regra de plural, no caso da saliência fônica; e de que a primeira posição, quando analisada isoladamente, favorece a aplicação da regra; assim como, a classe gramatical, cuja relevância se fez mais notável a partir da correlação com a variável posição ocupada pelo item lexical no SN. Além disso, fatores sociais como maior nível de escolaridade e o sexo masculino favoreceram significativamente a concordância nominal. No primeiro caso, os informantes de 5ª à 8ª série apresentaram peso relativo de 0,76, enquanto os de 1ª à 4ª e os analfabetos apresentaram, respectivamente, pesos de 0,46 e 0,28, confirmando tanto os estudos já discutidos até aqui quanto os demais que serão abordados mais adiante; no segundo caso, os homens tiveram um melhor desempenho no uso de marcas formais (0,60) em relação às mulheres (0,42), diferentemente dos dados encontrados por Scherre (1988) e Fernandes (1996). Uma possível explicação para isso seria, segundo a autora (1997), o fato da maioria dos homens trabalhar fora do núcleo social em que vive, estabelecendo um maior contato linguístico do que as mulheres que, na sua grande maioria, são donas de casa.

Outro trabalho de grande repercussão no âmbito da pesquisa sociolinguística brasileira foi o de Norma Lopes (2001), intitulado “Concordância nominal, contexto linguístico e sociedade”. Nesse estudo, a autora discute o fenômeno tomando por base aspectos atinentes ao processo de formação sócio-histórica da população de Salvador, justificando, para tanto, que essa cidade se caracteriza, sobretudo, por possuir

uma singularidade quanto a sua formação étnica: é uma cidade cuja população em maioria é negra, e foi palco, durante a época da colonização, de um enlace de culturas: foi sede do primeiro governo geral, no estado da Bahia aportaram os primeiros portugueses, nessa cidade habitaram grandes figuras da corte portuguesa e ela foi centro de grande concentração escrava urbana, e era de onde eles eram distribuídos para outras regiões, chegados através do porto de Salvador. Como as levas de escravos eram sucessivas durante muito tempo, sempre nela eram encontrados negros africanos aportados recentemente da África com suas diferentes línguas de origem. Dessa forma, tomou-se como hipótese a possibilidade de que, em Salvador tivesse havido por parte dos africanos um tipo de aquisição de língua com muita variação e que, mesmo sendo um contexto urbano, se pudesse, ainda hoje, identificar marcas dessa aquisição do português nos descendentes da população escrava (LOPES, 2001, P. 371).

Essa configuração étnica permitiu que a autora estabelecesse as seguintes variáveis sociais: sexo, dois grupos étnicos: um grupo de informantes com sobrenomes religiosos e outro de sobrenomes não religiosos, quatro faixas etárias: 15 a 24 anos, 25 a 30 anos, 45 a 55 anos e mais de 65 anos, além de três níveis de escolarização: 1 a 5 anos, 11 anos e mínimo de 15 anos de escolarização. A amostra foi extraída de inquéritos do Projeto Norma Urbana

Culta das décadas de 1970 e 1990, bem como do Programa de Estudos sobre o português Popular de Salvador. Daí, o propósito de Lopes em analisar as variedades culta e popular do português, mais precisamente no que se refere ao fenômeno da variação na concordância nominal de número.

No plano estritamente linguístico, a autora partiu de uma perspectiva atomística ou mórfica na qual foram destacadas como relevantes as variáveis: saliência fônica, marcas precedentes, classe gramatical associada à posição linear e posição relativa ao núcleo e contexto fonético-fonológico subsequente ao elemento do SN.

Os resultados obtidos com a saliência fônica revelaram que essa variável favorece a aquisição da regra de concordância, tendo em vista o princípio norteador de que as formas mais salientes, e por isso mesmo mais perceptíveis, são mais marcadas formalmente. A autora chegou a fazer o cruzamento dessa variável com a variável grau de escolaridade, o que resultou na óbvia constatação de que os informantes com maior nível de escolarização fazem muito mais a concordância formal de plural do que aqueles com menor nível.

Na análise da variável classe gramatical associada à posição linear e à posição relativa ao núcleo, Lopes (2001, p. 258) evidencia que os constituintes não nucleares antepostos ao núcleo recebem muito mais marcas formais de plural do que aqueles localizados à direita, sendo, portanto, marcados quase que categoricamente, com peso relativo de 0,81, numa frequência de 99% dos dados. Enquanto os pospostos ao núcleo foram menos marcados, numa frequência de 66% e peso relativo de 0,22. Esses resultados corroboram aqueles obtidos por Scherre (1988), Fernandes (1996) e Hebe Carvalho (1997), já discutidos no presente estudo.

Vale destacar também que o diferencial da análise de Lopes em relação às análises convencionais dos elementos flexionáveis do SN em função do núcleo é a adoção do princípio da adjacência, segundo o qual o constituinte sintagmático pode estar situado à esquerda e imediatamente seguido do núcleo ou à esquerda e não seguido de imediato ao núcleo. A autora conclui que o elemento à esquerda do núcleo e contíguo a este favorece bem mais a indicação de pluralidade do que o elemento não adjacente nessa posição. Em outras palavras, não é só o fato dos elementos se encontrarem antepostos ao núcleo que determina a maior ocorrência de marcação de plural, mas a proximidade do núcleo é o fator mais determinante, de acordo com a autora.

No que tange à variável marcas precedentes, os dados de Lopes dão conta de que a presença de numeral ou de qualquer outro elemento não marcado pelo morfema indicador de

plural antecedendo o constituinte de segunda posição, na estrutura sintagmática, não inibe a pluralização de tal constituinte. Ao contrário, acaba favorecendo.

Em se tratando do contexto fonético-fonológico subsequente, a autora deixa claro que o fator mais favorecedor de marcas de plural é a pausa final de sentença (0,62) e o que menos favorece é a consoante sonora (0,44).

Quanto às variáveis sociais analisadas pela autora, os resultados foram bastante conclusivos, permitindo, objetivamente, as seguintes constatações: quanto maior o grau de escolaridade maior a aplicação da regra de concordância no SN. Em relação à idade, a faixa etária mais velha, acima de 65 anos, foi a que mais favoreceu a marcação de plural (0,59). Com referência ao sexo, as mulheres fizeram mais uso de concordância do que os homens, fato que se modifica quando se considera na análise apenas os informantes com formação superior, de ambos os sexos. Nesse caso, houve nitidamente uma mesma frequência de uso padronizado da concordância.

O estudo de Elza da Silva Bueno (2002) sobre a influência de variáveis relacionadas à concordância de número envolveu 144 informantes de ambos os sexos, com pouca ou nenhuma escolaridade, residentes em distintas regiões do estado de São Paulo e que trabalhavam no corte de cana de açúcar. Ficou constatada na pesquisa da autora a forte influência das variáveis sociais no fenômeno de variação na concordância nominal. A autora chegou, inclusive, as seguintes conclusões: os informantes da faixa etária entre 15 e 29 anos foram os que mais usaram a concordância, sugerindo, com isso, segundo ela, uma provável incerteza dos jovens em relação ao futuro da profissão. As mulheres se apresentaram mais sensíveis à aplicação da regra e os falantes alfabetizados, embora com pouca escolarização, fizeram mais uso da concordância.

O trabalho de Odette de Souza Campos e Ângela Rodrigues (2002) se constitui em um estudo extremamente significativo, pois fornece dados relevantes acerca da concordância nominal na fala de informantes com formação em nível superior e com curso universitário em andamento. Para tal estudo, as autoras recorreram aos dados do projeto NURC, considerando os seguintes fatores: sociais, como idade, escolaridade e sexo; geográficos, com informantes de Porto Alegre, Rio de Janeiro, São Paulo, Salvador e Recife e estilísticos.

As autoras analisaram 3.011 SNs, explorando variáveis linguísticas como saliência fônica, posição do elemento no sintagma nominal e classe gramatical. Os dados de Campos e Rodrigues revelaram que a saliência não exerceu nenhuma influência na fala dos informantes, o que pode ser perfeitamente compreensível já que estes, em função do nível de escolarização, estão inseridos no contexto dos falantes da norma culta. Mas é importante lembrar, por outro

lado, que os estudos de Lopes (2001) e Hebe Carvalho (1997) evidenciaram a atuação dessa variável também na fala de informantes com nível superior.

A variável posição do elemento no SN foi considerada de forma isolada na análise das autoras, o que explica um maior índice de marcação de plural nos constituintes da primeira posição. Quanto à classe gramatical, essa variável não se mostrou relevante para a aplicação da regra de concordância nominal.

No que diz respeito às variáveis sociais, vale destacar, na pesquisa das autoras, a idade, uma vez que tanto os falantes da primeira faixa etária quanto os da terceira fizeram uso categórico das marcas formais de plural. Em relação ao sexo, as mulheres usaram mais o plural. A escolaridade se mostrou determinante para um percentual expressivo de ocorrências de pluralidade nos elementos do SN.

A variável estilística formalidade apenas confirmou o que as autoras já esperavam, ou seja, contextos de maior formalidade favorecem o uso das formas de prestígio, como a regra de plural no SN. Já a variável procedência geográfica apresentou resultados interessantes sobre o uso da concordância nominal entre os informantes das cinco capitais brasileiras envolvidos na pesquisa, deixando evidente que os de São Paulo aplicam mais a regra (0,97) enquanto os de Salvador (0,25) e de Porto Alegre (0,18) foram os que menos aplicaram.

O fenômeno da variação na concordância nominal também foi investigado por Patrícia Ribeiro de Andrade (2003) em uma comunidade rural de afrodescendentes, a de Helvécia (BA). A autora analisou a fala de 18 informantes, homens e mulheres, distribuídos em três faixas etárias: 21 a 40 anos, 41 a 60 anos, mais de 60 anos. Além de ter considerado as variáveis sociais: escolaridade e estado fora da comunidade.

As variáveis linguísticas foram analisadas sob duas perspectivas: sintagmática e mórfica. No plano mórfico, a autora definiu as seguintes variáveis: posição linear do constituinte, posição do constituinte com referência ao núcleo do SN, classe gramatical do constituinte, saliência fônica, marcas precedentes ao elemento nominal analisado e tonicidade. No plano sintagmático, foram analisadas as variáveis: configuração sintagmática do SN, função sintática do SN, posição do SN em relação ao verbo, número absoluto de constituintes do SN, número de constituintes flexionáveis do SN, saliência fônica dos elementos do SN e função textual do SN.

Na abordagem mórfica, merecem destaque no estudo de Andrade as variáveis: posição do constituinte no SN, marcas precedentes e saliência fônica. Em relação à primeira variável, os resultados foram bastante precisos ao confirmar a supremacia dos elementos que

antecedem o núcleo em termos de presença de marcas formais de plural, com a primeira posição sendo a mais marcada (0,94). A segunda variável acena para uma peculiaridade no dialeto de Helvécia, no que concerne ao elemento linguístico sob análise: a presença de marca formal precedendo tal elemento faz com que este tenha uma frequência muito baixa de pluralidade (0,45) e a ausência de marca no elemento anterior aumenta tal frequência (0,57). O resultado dessa análise rompe com o pressuposto defendido por Scherre (2008, p. 511) de que a língua, por meio da propriedade funcionalista, consegue agrupar formas semelhantes a partir dos processos mentais de associação, mais conhecido como processamento paralelo. Tal processamento, segundo ela, se apoia no princípio de que marcas levam a marcas e zeros levam a zeros. O que não foi, portanto, evidenciado por Andrade.

A saliência fônica foi outra variável utilizada na pesquisa da autora, e que se mostrou, num primeiro momento, irrelevante uma vez que os resultados envolvendo, individualmente, cada fator da escala acabaram contradizendo o princípio de que as formas mais salientes são mais marcadas do que as menos salientes. O que pode ser atestado pelos elevados índices percentuais de marcas de plural nos itens regulares (49%) e nos itens terminados em *-m* (63%), os menos salientes da escala. Mas quando os itens terminados em *-l*, *-r*, *-s* e os itens de plural duplo foram reunidos em um grupo e os ditongos nasais *-ão*: *-ães* e *-ões* em outro grupo de fatores, e considerando que os itens regulares compunham um terceiro grupo, percebeu-se que foi bem maior a probabilidade do primeiro grupo favorecer a inserção de marcas de plural (0,66), superando assim os regulares (0,47).

Em relação à abordagem sintagmática, é importante destacar que para a autora é “através da perspectiva de análise não atomística que podemos, efetivamente, nos certificar do nível de aplicação da regra de concordância de número entre os elementos do sintagma nominal” (ANDRADE, 2003, p. 104). E foi observando o comportamento das variáveis vinculadas a tal abordagem que a autora chegou à conclusão de que o uso da regra de concordância é muitíssimo baixo no dialeto de Helvécia. Os resultados gerais obtidos por ela na análise evidenciou essa tendência. Traduzindo os dados em números, visualiza-se melhor essa realidade linguística. Assim, no contexto de 1434 ocorrências de SNs, houve apenas 133 casos com concordância nominal, conforme previsto pela norma padrão. O que correspondeu a 9%, em termos percentuais, e a um peso relativo de 0,6. Mesmo reconhecendo os baixos níveis de concordância na comunidade, a autora julgou relevante observar o modo como se dá o encaixamento da variação na aplicação da regra de concordância nominal na estrutura linguística e social do dialeto, apoiando-se, para tanto, na análise de alguns grupos de fatores

linguísticos, como: configuração sintagmática do SN, saliência fônica dos elementos do SN, número de constituintes flexionáveis do SN e posição do SN em relação ao verbo.

De forma bastante resumida, é possível traçar o seguinte panorama do comportamento de cada uma dessas variáveis em relação ao fenômeno da variação de plural no SN. Na análise de Andrade (2003), observou-se que os sintagmas que mais favorecem a concordância nominal são aqueles formados por dois constituintes, sendo o primeiro um determinante e o segundo um numeral (*Ah, pra sê mais ô meno, pelo meno **uns três**, né?*), com frequência de 100%. Enquanto os mais inibidores são formados por estruturas do tipo: determinante + substantivo (***Umas** coisa que ele fala...mas **os médico de Texeira** é assim...*), com peso relativo de 0,46, e modificador + substantivo (*os **próprio remédio** era demais...*), com peso de 0,16. Ademais, a presença de um numeral mais um substantivo (*...que a gente tem **dois cano do osso**, né? Eu fui **duas veiz** em Tabuna faze perina.*) tende a desfavorecer a aplicação da regra formal, conforme verificado no peso de 0,62. O mesmo ocorre com a estrutura: Determinante + numeral + substantivo (*... quando deu **umas oito hora da noite**...*), de peso ainda menor (0,58).

A saliência fônica foi retomada pela autora na abordagem não atomística, distribuindo a escala de fatores em dois grupos: SNs apenas com itens regulares e SNs contendo itens que apresentam diferenciação de material fônico quando pluralizados. Com isso, obteve-se um peso relativo de 0,46 para o primeiro grupo e 0,66 para o segundo, confirmando assim a relevância dessa variável enquanto força condicionadora do uso das marcas formais de plural no SN.

Em relação ao número de constituintes flexionáveis no SN, ficou claro na análise da autora que os sintagmas mais favorecedores da regra de concordância são os que possuem apenas um constituinte flexionável (0,66). Já os sintagmas formados por três ou mais elementos pluralizáveis tendem a diminuir o número de marcas de plural (0,26).

No que diz respeito à posição do SN em relação ao verbo, os resultados de Andrade apontam para distintas tendências de aplicação da regra de concordância, na comunidade de fala de Helvécia, se comparado às inúmeras pesquisas já empreendidas nas quais essa variável tem sido analisada, como nos estudos de Scherre (1988), por exemplo.

Sobre as variáveis extralinguísticas ou sociais, as análises de Andrade evidenciam que em relação à faixa etária, tanto na perspectiva atomística quanto na perspectiva não atomística, a concordância de número no SN é mais utilizada pelos mais jovens (de 21 a 40 anos), enquanto entre os informantes mais velhos (acima de 60 anos) a marcação de plural é bem menos. Segundo a autora, esses resultados, aliados à sócio-história da comunidade,

sugerem que “o dialeto de Helvécia está passando por um processo de mudança em curso” (ANDRADE, 2003, p. 125). Isto, no que se refere ao fenômeno da variação na concordância nominal. A variável sexo, por sua vez, revelou que os homens tendem a liderar a mudança no sentido da implementação da regra de concordância, o que se justificaria pelo fato de eles estarem mais expostos às variantes de prestígio em função das atividades ocupacionais relacionadas a contextos econômico e social. A importância de tal exposição, em termos linguísticos, pode ser confirmada a partir da variável estada fora da comunidade cujo peso relativo foi de 0,61. Quanto à escolarização, a investigação da autora recaiu sobre a fala de pessoas analfabetas e semi-analfabetas, tanto na abordagem mórfica quanto sintagmática. Os resultados probabilísticos apontaram para uma tendência maior de uso da regra de plural entre os analfabetos. O que, em princípio, parece um tanto contraditório. Mas o certo é que para Andrade a “escola não exerce, na comunidade de Helvécia, influencia no que diz respeito ao desenvolvimento dialetal dos adultos” (ANDRADE, 2003, p. 134).

Além de todos esses estudos até aqui discutidos, alguns outros trabalhos mais recentes são referências imprescindíveis para se compreender com mais profundidade os mecanismos que regem o fenômeno da variação nos constituintes flexionáveis do SN. Nesse sentido, nos ocuparemos, igualmente, de pesquisas realizadas a partir de 2010 e que têm fornecido dados elucidativos acerca do caráter sistemático de tal variação, reforçando ainda mais a ideia de que o fenômeno em questão se configura, de fato, como caso típico de variação inerente, nos termos laboviano (LABOV, 2008, p. 183-259).

Assim, dando continuidade ao levantamento retrospectivo relacionado às pesquisas já realizadas sobre a concordância de número no SN, no português brasileiro, merece destaque o estudo de Lília Soares Miranda Santos (2010) que trata da ausência da concordância nominal de número entre falantes residentes em Pedro Leopoldo-MG. A autora estratificou a amostra dos 27 informantes, considerando grupo social, dividido em classe alta, média e baixa; idade, 17 a 23 anos, 40 a 47 anos e acima de 60 anos; sexo e grau de escolaridade: ensino fundamental, médio e superior. Mostraram-se relevantes na análise duas variáveis linguísticas: classe gramatical dos elementos não nucleares do SN e presença e ausência de flexão de marca plural nos elementos não nucleares do SN. Para uma melhor depuração dos resultados, Santos considerou a correlação sugerida por Scherre (1988) entre posição, classe gramatical e relação dos elementos não nucleares antepostos e pospostos ao núcleo, o que permitiu a ela chegar à conclusão, já esperada, de que os constituintes sintagmáticos antepostos ao núcleo são bem mais marcados do que os pospostos.

Já entre as variáveis extralinguísticas, o fator sexo mostrou que a ausência de marcas formais de plural é mais presente na fala dos homens. A escolaridade foi um fator que apenas ratificou o que a autora já previa e os estudos anteriores, por nós citados, já haviam constatado, ou seja, informantes com baixo nível de escolaridade tendem a usar muito menos as formas de prestígio. Os dados da variável grupo social, no que tange ao fator classe baixa, confirmaram também a hipótese de Santos segundo a qual a variante ausência de marcas formais de plural é muito mais utilizada pelos falantes das classes que se encontram na base da pirâmide social, com peso relativo, inclusive, de 0,67. Mas os resultados das análises dos fatores classe alta e classe média demonstraram uma realidade sociolinguística um tanto atípica, do ponto de vista da relação entre *status* social e uso das formas linguísticas de prestígio. Assim, a classe alta foi quem acabou favorecendo, surpreendentemente, a ausência das marcas formais de plural no sintagma nominal (0,45), enquanto a classe média favoreceu bem menos, conforme o peso de 0,33.

O trabalho de Sílvia Figueiredo Brandão (2011) envolveu também informantes com diferentes níveis de escolaridade: fundamental, médio e superior, sendo estratificados, além disso, em sexo e faixa etária. O *corpus* da pesquisa é composto por 18 entrevistados de Nova Iguaçu-RJ, embora a investigação da autora envolvesse também informantes de São Tomé e Príncipe. A análise das variáveis linguísticas revelou o que quase todas as pesquisas já demonstraram, ou seja, para a posição linear os constituintes que antecedem o núcleo são os mais marcados formalmente, superando tanto o núcleo quanto os elementos pospostos a este. A maior incidência de plural no constituinte nuclear ocorre quando este encontra-se na primeira posição. A animacidade, outra variável analisada pela autora, mostrou, conforme os resultados, que elementos sintagmáticos com traço [+animado], isto é, [+humano] favorecem mais a marcação de plural do que aqueles com traço [-animado]. Em relação à saliência fônica, os resultados de Brandão foram bastante conclusivos ao constatar que quanto mais perceptível for um item lexical na oposição singular/plural mais marcas de plural ele tende a receber.

Para as variáveis sociais, os dados da autora deram conta de que os jovens, no fator faixa etária, é quem mais fazem uso das marcas de plural. O fator grau de escolarização demonstrou o que já é quase consenso entre os estudiosos: quanto maior o nível de escolaridade do falante maior o índice de marcas de plural.

Em sua pesquisa, Janaína Biancardi da Silva (2011) investigou também a fala de informantes com diferentes níveis de escolarização, distribuídos em faixas etárias e sexo. O *corpus* utilizado pela autora foi o do Projeto Português Falado da cidade de Vitória-ES e

envolveu 43 entrevistas. Os dados relacionados à concordância nominal de número foram analisados considerando tanto esses fatores sociais quanto os linguísticos: posição do elemento em relação ao núcleo e saliência fônica. O resultado probabilístico dessa primeira variável (0,81) reforça ainda mais as constatações da grande maioria dos estudos apresentados nesse capítulo de que os elementos pré-nucleares exercem muito mais influência na marcação de plural no SN do que aqueles que se localizam à direita do núcleo, cujo peso relativo ficou em 0,15. A mesma força condicionadora das marcas de plural teve a saliência fônica. A análise da autora demonstrou que os itens lexicais mais salientes receberam mais marcas formais de plural (0,61) do que os menos salientes (0,47).

No que diz respeito às variáveis sociais, observou-se que, entre as faixas etárias selecionadas, os jovens entre 15 e 25 anos foram os que mais fizeram uso das marcas de plural no sintagma nominal (0,61). Quanto ao sexo, os homens foram os mais favorecedores (0,58). A escolarização seguiu o mesmo padrão previsto: quanto maior o nível, maior o uso das formas prestigiadas pela norma culta. O que se comprovou com os seguintes resultados: peso relativo de 0,65 entre os informantes universitários, 0,54 entre os informantes do ensino médio e 0,32 entre os informantes do ensino fundamental.

De fato, o acesso à cultura letrada é um fator preponderante para a aquisição e uso das formas linguísticas de prestígio. E a depender do nível de educação formal em que se encontram os informantes, é possível até coletar dados de fala com altos índices de marcação de plural ao longo do SN. Foi o que aconteceu na pesquisa de Flávio Marcelo Bueno de Castro e Vinícius Carvalho Pereira (2012) que analisaram dados de 30 pessoas com nível universitário da cidade de Cuiabá-MT, obtendo assim uma frequência em torno de 96% de uso de marcas formais de plural. No plano linguístico ou estrutural, os autores elegeram apenas a variável *distribuição dos elementos no SN de acordo com a classe gramatical*. E os números revelaram que os determinantes foram os mais marcados, com percentual de 84% contra 33% do núcleo. O que não é muito diferente das inúmeras pesquisas aqui discutidas.

O trabalho realizado por Simone Daise Schneider (2012) se constitui em um estudo bastante singular envolvendo a concordância nominal. A autora investigou a fala de 30 crianças de 3 a 6 anos de idade da cidade de Novo Hamburgo-RS. Oriundas de famílias de nível escolar médio e superior e de classe social média-alta, essas crianças, embora ainda não alfabetizadas, já cursavam o ensino infantil. Nos dados analisados pela autora, em torno de 3028 itens flexionáveis, 95% deles receberam as marcas formais da flexão nominal de número. Um resultado que surpreende, tendo em vista que em quase todas as pesquisas apresentadas e discutidas por nós nesse nosso estudo vinculam o maior uso da regra de

concordância nominal a um nível maior de escolaridade. Mas é preciso lembrar, por outro lado, que essa vinculação tem sido associada às falas de informantes adultos uma vez que pesquisas variacionistas envolvendo o público infantil parecem um tanto incipientes no Brasil. Além disso, a autora lembra que

[...] os dados desta pesquisa são, basicamente, oriundos de sintagmas nominais plurais formados por dois elementos (determinante + nome), no máximo, três elementos (determinante + nome + adjetivo). Esse fato, certamente, contribuiu para o alto índice de aplicação da regra, porque, segundo a literatura, exatamente nessas posições que as marcas de plural estão mais presentes (SCHNEIDER, 2012, P. 135).

Ora, em se tratando de crianças, sobretudo a faixa etária de 3 a 6 anos, deve-se considerar o fato de que as mesmas estão em fase de aquisição de estruturas linguísticas mais elaboradas, como a morfossintaxe, por exemplo. E nesse sentido, o contato social estabelecido no contexto familiar – todas são de famílias de classe média-alta, sendo que a maioria dos familiares tem formação em nível superior - e na própria escola se torna fundamental para o processo de assimilação e incorporação de hábitos linguísticos mais prestigiados, socialmente. Schneider (2012, p. 136) ainda justifica que “as crianças aqui pesquisadas caracterizaram-se por estarem aprendendo uma variedade do português falado no Brasil que tende a marcar a flexão de plural na maior parte de suas produções”.

Vejamos, pois, as variáveis linguísticas e extralinguísticas analisadas por Schneider a fim de percebermos melhor o comportamento da variável dependente, concordância nominal de número, na fala dos informantes infantis. Apresentaremos os resultados de algumas variáveis que se mostraram bastante relevantes na análise da autora, lembrando que ela considerou, em seu estudo, abordagens atomística e não atomística.

Na perspectiva não atomística, foram analisados 1626 SNs, constituídos, na sua maioria, de dois elementos. 91% desses sintagmas receberam marcas formais de plural em todos os constituintes. A faixa etária de 5,1 a 6 anos foi a que mais favoreceu a indicação de pluralidade (0,59). As meninas tiveram um desempenho mais satisfatório (0,59) do que os meninos (0,46).

Em relação à abordagem atomística, merecem destaque na pesquisa da autora as seguintes variáveis independentes, com seus respectivos resultados:

- Processos morfofonológicos de formação de plural, nos quais foram considerados três fatores: itens com plural metafônico (0,22), item regular (0,50) e itens com modificação da sílaba final (0,56);

- Tonicidade dos itens lexicais: proparoxítona (0,63), oxítono e monossílabo tônico (0,57) e paroxítona e monossílabo átono (0,48);
- Posição linear dos elementos no sintagma nominal: 1ª posição (0,72), 2ª posição (0,31) e 3ª posição (0,35);
- Classe gramatical dos elementos: determinante(0,71), categoria substantivada (0,62), substantivo (0,30) e adjetivo (0,19);
- Contexto fonológico seguinte: item seguido de vogal(0,71), item seguido de consoante (0,55) e item seguido de pausa (0,34);
- Grau dos itens lexicais: normal (0,51) e diminutivo/aumentativo (0,32);
- Posição do elemento em relação ao núcleo do sintagma nominal: elemento anteposto ao núcleo (0,70), elemento é o próprio núcleo (0,31) e elemento posposto ao núcleo (0,22).
- Faixa etária: 5,1 a 6 anos (0,55), 3 a 4 anos (0,47) e 4,1 a 5 anos (0,45);
- Sexo: meninas (0,53) e meninos (0,46);
- Interação: narrativa tradicional (0,71), zoológico (0,57), lojinha (0,56), fazendinha (0,40), reconto (0,39) e narrativa espontânea (0,38).

No que se refere ao plural metafônico, é importante que se considere que o baixo índice de marcas de plural nesse fator, sobretudo entre as crianças de até 5 anos, se deve, segundo a autora, ao fato delas estarem passando por um processo de aprendizagem fonológica. O que seria um argumento favorável para a constatação de que esse seria um fator, no eixo da saliência fônica, em sua dimensão processo, um tanto desfavorecedor da aplicação da regra de plural no *corpus*. Em relação à tonicidade dos itens lexicais, Schneider justifica o motivo de a categoria proparoxítona ter sido a que mais favoreceu a marca morfológica de plural, alegando que o número de palavras foi muito pequeno, apenas 7. Quanto à variável sexo, a autora chegou a fazer o cruzamento com faixa etária, o que evidenciou ainda mais um acentuado uso de marcas de plural na fala das meninas. Os fatores selecionados para a variável interação, por sua vez, permitiram que a autora observasse, na análise, a marcação de plural no SN na produção oral infantil em diferentes contextos interativos.

A pesquisa de Vívian Antonino (2012) envolveu também a concordância nominal, porém, a autora analisou o fenômeno da variação em predicativos do sujeito e em estruturas passivas, no português popular de Salvador-BA. O índice geral obtido por ela foi de 14,6% de aplicação da regra de plural. Para chegar a tal resultado, a autora considerou algumas

variáveis independentes que se mostraram estatisticamente relevantes e que julgamos oportuno apresentá-las aqui, ainda que de forma resumida, conforme relação abaixo:

- Material interveniente entre o verbo e o predicativo: a concordância nominal de número é desfavorecida se houver qualquer material, com percentagem de 1,7% e peso relativo de 0,264. Ao passo que a ausência de material favorece a concordância (18,8% / 0,736);
- Concordância nominal de número no SN sujeito: sintagma nominal com concordância formal favorece a concordância de número nos predicativos e nas passivas (34,8% / 0,780), já os sintagmas nominais sem concordância formal desfavorecem (3,6% / 0,220);
- Concordância verbal: sentença com concordância verbal favoreceu as marcas de plural em predicativos e passivas (29,2% / 0,750), enquanto a ausência de concordância verbal acabou desfavorecendo (3,0% / 0,250);
- Tipo de atributo: atributos essenciais, ligados pelo verbo ser, favoreceram a marcação de número (20,6% / 0,772). Os atributos acidentais desfavoreceram (1,4% / 0,228);

Além dessas variáveis, Antonino analisou também: tipo de predicação, na qual ela constatou haver maior indicação de pluralidade em predicativos do sujeito do que em estruturas passivas; animacidade do sujeito, cujo traço [+ humano] favoreceu a aplicação da regra de número, ao contrário do traço [- humano] que contribuiu para o desfavorecimento; saliência fônica, que confirmou a hipótese de que os itens irregulares, por se mostrarem mais salientes, são mais marcados do que os regulares. A autora ressalta, igualmente, a variável nível de exposição à mídia, destacando o fato de que aqueles falantes que se expõem mais à mídia fizeram mais a concordância nominal do que aqueles que afirmaram ter pouco contato. Houve, inclusive, um nível de marcação de plural elevado entre informantes mais expostos a programas religiosos.

Flávia Santos Martins (2013) investigou a variação na concordância nominal de número na fala dos habitantes de cinco localidades pertencentes à microrregião do alto Solimões (Amazonas). Foram 57 informantes, homens e mulheres, envolvidos na pesquisa, distribuídos em três faixas etárias: 18 a 35 anos, 36 a 55 anos e acima de 56 anos. Os resultados dos dados analisados pela autora dão conta de que o maior índice do uso formal de plural foi encontrado entre os falantes da primeira faixa (61% / 0,55), o que sugeriria, de acordo com ela, uma possível mudança em progresso em direção a um sistema com a presença de marcas formais/informais de plural. As mulheres se apresentaram mais sensíveis

à aplicação da regra (60% / 0,52) do que os homens (56% / 0,47). A escolaridade foi outro fator que confirmou a conjectura estabelecida pela autora segundo a qual informantes com maior tempo e nível de escolaridade fazem muito mais uso das formas gramaticais mais prestigiadas socialmente, como é o caso da concordância nominal entre os elementos do SN. A variável foi analisada considerando dois fatores: 9 a 11 anos (64% / 0,57) e 4 a 8 anos (54% / 0,43). Dentre algumas outras variáveis sociais observadas por Martins, destacamos a *ocupação*. Os dados demonstraram que os informantes da categoria alta: estudante, professor, professor aposentado, funcionário público, gestor, funcionário público saúde, agente de saúde, auxiliar de contabilidade, vereador e auxiliar administrativo tiveram melhor desempenho na aplicação da regra de concordância nominal, com frequência de 63% e peso probabilístico de 0,55. Resultado levemente superior à categoria média (60% / 0,53), composta por: comerciante, autônomo, entregador, moto-táxi, fazendeiro e atendente. Enquanto o índice de plural no SN entre os falantes da categoria baixa - agricultor, carregador, dona de casa, auxiliar de serviços gerais, pescador, manicure, cozinheiro e aposentado - foi bem menor (53% / 0,43). Essa variável permitiu que a autora atestasse a hipótese por ela levantada de que

[...] os informantes cuja ocupação no mercado de trabalho “exige” formas linguísticas prestigiadas socialmente são os favorecedores da aplicação da regra (alta: 0,55; média: 0,53), enquanto os informantes que estão envolvidos em ocupações que não “exijam” o uso da variante prestigiada socialmente desfavorecem a aplicação da regra (0,43) (MARTINS, 2013, p. 175).

No âmbito das variáveis estruturais, também conhecidas como variáveis linguísticas, a autora observou: a posição dos elementos não nucleares em relação ao núcleo/núcleo, cujos resultados confirmaram muitos dos estudos por nós aqui discutidos, ou seja, elementos não nucleares antepostos ao núcleo são os que mais recebem indicação de pluralidade, conforme frequência de 95% e peso relativo de 0,78; cruzamento entre posição dos elementos não nucleares em relação ao núcleo/núcleo e posição linear, em que o objetivo da autora foi verificar o comportamento diferenciado do núcleo de acordo com a posição ocupada no SN. O que permitiu que ela chegasse aos seguintes dados percentuais: elementos não nucleares antepostos: 1ª posição (96%), 2ª posição (91%) e 3ª posição (57%); elementos não nucleares pospostos: 2ª posição (45%) e 3ª posição (24%); núcleo: 1ª posição (96%), 2ª posição (30%), 3ª posição (31%) e 4ª posição (29%). As duas dimensões da saliência fônica correlacionadas: processos morfofonológicos de formação de plural e tonicidade dos itens lexicais forneceram os resultados a seguir encontrados na amostra da autora, segundo o grau da saliência (+ / - diferenciação fônica na relação singular e plural): plural com alternância vocálica nos itens

terminados em -l (77% / 0,90), plural nos itens terminados em -ão, com alteração silábica (75% / 0,86), plural nos itens terminados em -s (74% / 0,86), plural nos itens terminados em -r (70% / 0,85) e plural duplo com alternância vocálica (55% / 0,75). Todos esses fatores foram considerados mais salientes. Por outro lado, se apresentaram menos salientes os fatores: nomes regulares oxítonos e monossílabos tônicos (72% / 0,62), regulares paroxítonos (42% / 0,44) e regulares proparoxítonos (25% / 0,37). Os resultados dessa variável confirmaram a pressuposição de Martins (2013, p. 147) de que “itens com formação de plural irregular são os que mais favorecem a aplicação da regra, enquanto os de formação regular as desfavorecem”.

Outras variáveis linguísticas que se mostraram relevantes na pesquisa da autora foram: *marcas precedentes*, que comprovou a hipótese de que a presença de marca formal de plural anterior ao elemento em análise leva a presença de marca em tal elemento. Na verdade, é o princípio de que marcas levam a marcas e zeros levam a zeros, discutido por Scherre (1988) que se apoiou no Processamento com Paralelismo, segundo o qual há uma tendência na língua das formas semelhantes se agruparem. Já a análise do *contexto fonético-fonológico subsequente* acabou não confirmando a hipótese da autora de que a pausa seria o fator mais favorecedor da presença de plural. Os dados revelaram que a presença de uma vogal é que mais favorece a regra (66% / 0,59), seguidas de pausa (36% / 0,55) e consoante (66% / 0,45). E, finalmente, vale ressaltar também os resultados encontrados pela autora em relação à variável *características dos itens lexicais (substantivos e adjetivos)*. Na análise, ficou evidente que o grau aumentativo tende a ser o mais marcado (40% / 0,68) e o diminutivo o que recebe menos marcas formais de plural (14% / 0,32).

Maria Aparecida Souza Guimarães²⁶ (2014) examinou o vernáculo de falantes de Vitória da Conquista, considerando a variação de número no SN. A investigação provou que fatores estruturais ou linguísticos condicionam a aplicação da regra de concordância no SN: posição linear do constituinte, posição do constituinte com referência ao núcleo do SN, classe gramatical do constituinte e saliência fônica. Em relação às variáveis sociais, pudemos perceber que as mulheres e os mais jovens tendem a aplicar mais a regra de concordância. A tendência de aplicação da regra verificada na população de Faixa I no *corpus* demonstra que a mudança em curso é um processo decorrente da ampliação de horizontes desse grupo, pois ele está mais em contato com os grupos externos e sente a pressão que deles procede no sentido de se adaptarem à vida urbana. Além disso, eles são os mais atingidos pelos meios de comunicação, pela televisão e pelo rádio, que imprimem valores diferentes dos estabelecidos pelo ambiente familiar ou pelo círculo de relações.

²⁶ O estudo de Guimarães (2014) será posteriormente discutido para fins de comparação com o presente estudo.

Como se pode notar, as pesquisas sobre a variação na concordância nominal de número no SN já cobrem um número significativo de trabalhos realizados em território brasileiro, todos empreendidos à luz da teoria variacionista de cunho laboviano. Os dados de fala coletados em várias partes do país sinalizam para uma certa uniformidade do fenômeno no sentido de que ele ocorre nos mais diversos contextos geográficos e sociais, porém com frequência flutuante, conforme a região e o *status* social dos falantes. Os pesquisadores têm adotado em suas análises variáveis linguísticas e extralinguísticas cujos resultados evidenciam o caráter inerente da variação no sistema flexional de número no português brasileiro. Todos os estudos que mencionamos nesse capítulo já seriam suficientes para mapear as localidades onde a variável tem sido investigada, bem como, a frequência de uso da variante *presença de marcas formais de plural*. Mas, por ora, esperamos que esse nosso estudo possa contribuir para o conhecimento da realidade sociolinguística da comunidade de fala de Vitória da Conquista, considerando, sobretudo, a necessidade do município se firmar como mais uma área dialetal brasileira com um fecundo campo sociolinguístico a ser cada vez mais explorado.

Nossa pesquisa se apoiará nos estudos aqui apresentados, tendo em vista que os resultados das análises obtidos pelos autores se mostraram bastante conclusivos do ponto de vista estrutural e social. Discorreremos, com mais exaustão, sobre o trabalho de Guimarães (2014), pois é a partir deste que iremos analisar e discutir, comparativamente, a variação na concordância nominal de número no português popular e culto de Vitória da Conquista. No capítulo a seguir, apresentamos os procedimentos metodológicos de nossa pesquisa.

3 METODOLOGIA

No presente capítulo, seguindo a metodologia proposta pela Sociolinguística Variacionista de cunho laboviana, apresentamos os procedimentos metodológicos norteadores de nossa pesquisa. Assim, descrevemos tanto como foi realizada a pesquisa de campo quanto os critérios adotados para a elaboração das hipóteses explicativas responsáveis pela escolha das variáveis linguísticas e extralinguísticas controladas. Devemos destacar que, em nosso estudo que enseja a comparação entre as normas popular e culta, utilizaremos amostras de fala de 24 (vinte e quatro) informantes, sendo 12 (doze) para cada norma.

Nosso objetivo principal é descortinar tendências de aplicação da regra de concordância de número no Sintagma Nominal, estabelecendo uma comparação entre as normas popular e culta, tomando como *corpus* amostra de falas de informantes naturais de Vitória da Conquista – BA. Assim, adotamos os pressupostos teóricos e os procedimentos metodológicos fornecidos pela Teoria da Variação e Mudança Linguística propostos por Weinreich, Labov e Herzog (2006), Labov (2008), em consonância com os fundamentos da sócio-história (LUCCHESI, 2000). Como apontou Guimarães (2014), a variação na concordância nominal é um fenômeno que pode ser explicado por influxo de fatores sociais e linguísticos, controláveis a partir de recortes teóricos. Em nosso caso, lançaremos mão de uma análise atomística tal qual proposta de Scherre (1988) ²⁷.

3.1 PROCEDIMENTOS METODOLÓGICOS DA PESQUISA

Com o florescimento das pesquisas *in loco* entre os séculos XVIII e XIX, as Ciências Sociais se viram em meio a um paradoxo no sentido de coletar dados sem que houvesse a “manipulação” do pesquisador. Destarte, os fenômenos observados deveriam reproduzir um cenário de normalidade verificável sem a presença de um agente externo. Marconi e Lakatos (1990) asseveram que a pesquisa de campo pode restringir-se ao simples observar, mas pode incluir o contato entre o agente interno (o pesquisado) e o externo (o pesquisador) por meio da interação da entrevista. Assim, dentro do aparato metodológico da Sociolinguística laboviana, realizamos entrevistas, as quais são precedidas de um contato inicial para que possamos fazer, por meio de questionário, um levantamento social a fim de verificarmos se o informante enquadra-se no perfil desejado pelo recorte da pesquisa. Em nosso caso, os critérios (para se

²⁷ Scherre realiza em 1988 duas análises: a sintagmática e a atomística. Em nosso estudo, apenas priorizamos a última forma de análise.

constituírem informantes) foram: local de nascimento/procedência, escolaridade, sexo e idade. Ainda compondo o quadro metodológico de nossa pesquisa, fizemos um levantamento da formação sócio-histórica da comunidade de fala de Vitória da Conquista, alvo de nossa investigação sociolinguística - que apresentaremos logo a seguir - a fim de conhecermos melhor a realidade da região. Na sequência, limitamo-nos a descrever cada etapa da pesquisa realizada na localidade em questão.

3.1.1 Perfil sócio-histórico da comunidade de fala de Vitória da Conquista

Em seu trabalho intitulado *História de Conquista: Crônica de uma cidade*, Mozart Tanajura (1992, p. 17) apresenta de forma bastante objetiva a localização espacial dessa cidade cujos limites fronteiriços apontam para o Norte, limitando-se com os municípios de Anagé e Planalto; o Sul, com Encruzilhada e Cândido Sales; o Leste, com Barra do Choça e Itambé e o Oeste, com os municípios de Belo Campo e Anagé. A posição geográfica em que se encontra a cidade de Vitória da Conquista, localizada na região sudoeste do estado da Bahia, mais precisamente na microrregião do Planalto de Conquista, é um indicativo de que o município está situado em uma área estratégica que lhe permite estabelecer relações comerciais com diferentes localidades. Ainda, segundo o autor, até 1950 a área do município era de aproximadamente 9.422 Km², porém, com a emancipação de alguns distritos, a área foi diminuída para 3.743 Km².

A ocupação da região de Vitória da Conquista esteve relacionada, segundo Sousa (2013, p. 26) “a interesses de ordem privada”. Esses interesses, para ela, envolvia a busca por terras e riquezas minerais, sobretudo, na região conhecida como Sertão da Ressaca²⁸. O pioneirismo desbravador se deve, segundo Silva-de-Sousa (1998, p. 35), a João Gonçalves da Costa que havia se fixado na região a partir de 1744, tornando-se um grande proprietário de terras e grande criador de gado. Atribui-se a ele a fundação do Arraial da Conquista por volta de 1752, como observa Sousa (2013, p. 26).

No contexto político, Silva-de-Sousa (1998) destaca o fato de que, após mudanças de ordem administrativa e territorial, o arraial se elevou à condição de Imperial Vila da Vitória, em 1840 e em 1891 passou à categoria de cidade, recebendo o nome de Conquista, e posteriormente, em 1943, o nome do Município é modificado para Vitória da Conquista.

²⁸ Para Novais (2008), Sertão da Ressaca refere-se à região do sudoeste da Bahia, onde se localiza a cidade de Vitória da Conquista. Medeiros (citado por SILVA-DE-SOUSA, 1998, p. 10) lembra que o termo era amplamente aplicado às terras situadas entre os rios Pardo e das Contas.

A emancipação e consolidação do poder político fez de Vitória da Conquista um grande centro de integração comercial, apresentando um desenvolvimento econômico privilegiado, sobretudo, no setor de produção agropecuária, como destaca Fontes (2009, p. 03). Porém, foi no final da década de 1940 e início da década de 1950 que Vitória da Conquista entrou de vez na rota do desenvolvimento e do progresso graças às políticas públicas de construção de rodovias patrocinadas pelos governos federal e estadual, como lembra Sousa (2013, p. 28). A partir daí, a cidade se tornou um entreposto comercial de destaque no sertão da Bahia, viabilizando a ligação com outros centros econômicos do estado e do país.

Esses acontecimentos vão repercutir de forma bastante pujante na dinâmica populacional e social da região. A migração cada vez maior da zona rural para a sede do município, sem contar os migrantes oriundos de outras partes do estado e até mesmo do país, mudou radicalmente a realidade demográfica do município num intervalo de duas décadas, conforme Tabela 3, apresentada por Angelita Cunha da Silva Sousa em sua dissertação de mestrado defendida em 2013, intitulada “A Rua do Maga-Sapo: cotidiano e representações da prostituição em Vitória da Conquista-BA (1950-1971)”:

Tabela 3 – Distribuição da população rural e urbana do município de Vitória da Conquista (1940 a 1970)

Ano	População rural	%	População urbana	%	Total
1940	24.910	74,3%	8.644	25,7%	33.554
1950	26.993	58,4%	19.463	41,6%	46.456
1960	31.401	39,3%	48.712	60,7%	80.113
1970	41.569	32,5%	85.959	67,5%	127.528

O vertiginoso crescimento populacional urbano demonstrado na **tabela 3**²⁹ reflete justamente as mudanças que estavam ocorrendo na cidade em função, principalmente, da abertura da BR -116. Portanto, a urbanização acelerada e o desenvolvimento econômico, ambos incrementados com a implantação da rodovia, se constituíram em dois vetores importantes para se compreender a composição social e demográfica da região na segunda metade do século XX, com a configuração praticamente de uma inversão de valores com relação à População Rural (PR) e à População Urbana (PU) no intervalo de 30 anos, entre as

²⁹ Na verdade, essa tabela encontra-se, originalmente, publicada no jornal **Fifó**, do dia 11 de outubro de 1997, na página 9. A matéria é de autoria de Ruy Medeiros e traz como título “Aspecto urbano de Conquista através da História”.

décadas de 40 e 70, Além do mais, é importante pensar no fenômeno migratório ocorrido em Vitória da Conquista, em que a população urbana superou a população rural, como evento caracterizador da própria realidade migratória brasileira que já apresentava, na passagem da década de 1960 para 1970, um contingente populacional muito maior na zona urbana, conforme dados do IBGE.

Nada mais justo, nesse momento, do que resgatar a história daqueles que foram, de fato, os legítimos proprietários das terras onde se encontra localizado o município de Vitória da Conquista: os povos indígenas. O trabalho de Silva-de-Sousa (1988) se constitui em uma fonte bastante esclarecedora no que tange aos aborígenes que habitaram a região. Ela destaca a presença de três grupos: os pataxós, os mongoiós e os aimorés³⁰.

Esses povos foram duramente perseguidos por João Gonçalves da Costa, um legítimo bandeirante que se instalou no sertão da ressaca. As suas ações no combate aos nativos renderam-lhe muito prestígio perante as autoridades da capitania. Como bem lembram Silva e Sousa (2013), “pouco restou da tradição e dos costumes indígenas no Sertão da Ressaca”, destacando tão somente a tradição da fabricação de objetos de barro cozido, em alguns bairros periféricos, “como influência direta da cultura indígena que foi dizimada”.

No plano linguístico, julga-se pertinente reproduzir aqui o quadro ilustrativo utilizado por Silva e Sousa, contendo as línguas faladas pelos povos indígenas que habitaram a região de Vitória da Conquista, conforme apresentação abaixo:

Quadro 2: Línguas faladas pelos indígenas que habitaram a região de Vitória da Conquista

Família	Língua	Dialeto
Kamacã	Kamacã	Mongoyó
Aymoré	Aymoré	-----
Maxacali	Pataxó	-----

Além dos indígenas, os negros africanos e afrodescendentes influenciaram também a cultura e a formação sóciolinguística da região. Alguns documentos, como os inventários, por exemplo, têm se constituído como fontes seguras para se certificar da presença da mão-de-obra escrava nas fazendas do município, sobretudo no período que antecedeu a abolição da escravidão no Brasil.

³⁰ São conhecidos também como *imborés* ou *botocudos*. Segundo o príncipe Maximiliano, “o nome de “botocudos” lhes vem de usarem, no lábio e nas orelhas, uma grande cavilha de madeira, à semelhança de batoque, que é como chamam os portugueses as rolhas de barril” (WIED-NEWIED, 1940, p. 274).

Os estudos realizados por Idelma Aparecida F. Novais, relacionados à produção e comércio na Imperial Vila da Vitória, apresentam dados relevantes sobre o trabalho escravo utilizado entre os anos de 1840 e 1888. Baseado num número significativo de inventários consultados, 207 ao todo, a autora fez uma pesquisa bastante reveladora no que diz respeito ao número de inventariados por faixa de escravos e o percentual que os cativos tinham na formação das fortunas dos seus proprietários.

Nos inventários, consta um contingente aproximado de 1.260 escravos compondo a organização populacional da então Imperial Vila da Vitória. Mas no livro de matrícula³¹ da vila, de 1871, Novais (2008) encontrou um número maior de escravos matriculados, 1590, Santos e Nascimento, (2010), apoiando-se no trabalho de Ivo (2004)³², apresentam uma quantidade ainda maior de escravos, 1.846, conforme os dados do censo de 1870.

Independentemente do número exato, essa quantidade de negros escravizados representou uma população relevante numa região que estava passando por um processo de povoamento. Segundo Nascimento (2008 *apud* Santos e Nascimento, 2010), os escravos da Imperial Vila da Vitória, no final do século XIX, eram quase todos nascidos no Brasil e havia um número equilibrado entre homens e mulheres. Já nos inventários pesquisados por Novais (2008), pertencentes aos maiores proprietários de escravos, em apenas um – o de Lourenço Gonçalves da Costa e Ana Senhorinha de Jesus – consta a origem dos cativos. Na descrição feita pela autora, há escravos de várias regiões da África, tais como: “sete de Angola, um do Congo, um Aussá, um de Moçambique” (NOVAIS, 2008, p. 57). O que, linguisticamente, sugere a convivência, em um mesmo espaço, de diferentes variedades dialetais.

Já a presença do homem branco, europeu, é sintetizada, ironicamente, na figura do conquistador João Gonçalves da Costa que, na verdade, era um ex-escravo nascido em Portugal, mais precisamente na cidade de Chaves, conforme pode ser verificado no trabalho de Silva-de-Sousa (1998).

3.1.1.1 O acesso ao letramento na comunidade de fala de Vitória da Conquista

Para discutirmos a história da educação em Vitória da Conquista, ainda que em sua fase primeva, cremos ser necessário inseri-la no contexto mais amplo da educação brasileira

³¹ Por meio de consulta ao Arquivo Público da Bahia (APEB) – Salvador-BA, a autora teve acesso a tal livro. A fonte completa é: APEB. Seção Colonial e Provincial. Série Correspondências. Março 1464. 1860-1889.

³² Para mais detalhes sobre a população negra de Vitória da Conquista no final do século XIX, consultar IVO, Isnara Pereira. *O Anjo da Morte contra O Santo Lenho: poder, vingança e cotidiano no sertão da Bahia*. Vitória da Conquista: Edições UESB, 2004.

de então. A esse respeito, são oportunas as palavras de Magalhães (2006), que ao discutir a educação tanto no plano regional quanto local, considera que “as regras de constituição das escolas e de suas transformações devem estar em função do panorama social que a circunda ou a circundou, em sua conexão com a política educacional e conjuntura da época” (MAGALHÃES, 2006, p. 3).

Como o nosso interesse recai sobre o ensino em Vitória da Conquista, no período que coincide com as primeiras décadas do período republicano, julgamos pertinente traçar, ainda que de forma bastante resumida, o panorama educacional do Brasil com o advento da república. Destarte, parece consentâneo trazer à baila a afirmação de Souza (2009), segundo a qual “os republicanos tomaram a educação como uma condição fundamental para o desenvolvimento da nação, buscando desenvolvê-la por todo o país, através de campanha em prol da alfabetização, criação de escolas públicas etc.” (SOUZA, 2009, P. 47).

Ora, embora dos ideais republicanos emergissem a imperiosa necessidade da obrigatoriedade do ensino, sobretudo o ensino laico, como condição imprescindível para o desenvolvimento da nação, verificou-se que nos primeiros anos republicanos em Vitória da conquista o analfabetismo, conforme a autora, “era predominante [...], correspondendo às expectativas do paternalismo dos coronéis que não se manifestavam muito a favor da escolarização da população, sobretudo a mais carente” (SOUZA, 2009, P. 49). A autora lembra ainda que das poucas escolas que ofereciam o ensino no município, no período correspondente à primeira república (1889-1930), a maioria delas funcionava na residência do professor. Essa situação começou a se modificar a partir da década de 1930 quando

[...] foi fundada a primeira escola estadual pública – o prédio escolar Barão de Macaúbas, fundado em 1935, que possuía apenas o curso primário -, o primeiro ginásio – o Ginásio de Conquista, primeiro estabelecimento de ensino ginásial³³ da cidade, de caráter particular, que começou a funcionar em 1940 – e a primeira Escola Normal (SOUZA, 2009, P. 51).

No plano didático-pedagógico, envolvendo o ensino da língua portuguesa, um levantamento exaustivo que fizemos, no Museu Pedagógico de Vitória da Conquista, sobre a bibliografia utilizada na época, permitiu-nos elencar os seguintes títulos: *O que se não deve dizer*; *Lições Práticas da Língua Portuguesa*; *Gramática Sintática da Língua Portuguesa* e *Novas Reflexões sobre a Língua Portuguesa*, todas de Cândido de Figueiredo. Além de *Língua Portuguesa, Gramática e Exercícios*, de Aníbal Bruno; *Curso de vernáculo*, de Cláudio

³³ Em seu estudo sobre *O Ginásio de Conquista – A memória de uma instituição escolar (1940-1960)*, Oliveira (2009, p. 74) lembra que o Ginásio foi fundado e dirigido pelo padre Luiz Soares Palmeira.

Brandão; *O idioma nacional*, de Antenor Nascentes e *Primeiro Livro de Leitura*, de Felisberto de Carvalho, que chegou a atingir a 119.^a edição.

Portanto, o acesso ao letramento passava necessariamente pela aprendizagem da norma padrão. Havia uma preocupação exacerbada em ensinar os padrões linguísticos europeus, tais quais prescritos pelas gramáticas. Essa situação por si só já é um indicativo do quanto a apropriação das formas de prestígio se constituía em um divisor de águas no contexto linguístico comportamental, tendo em vista que a língua reflete a hierarquia social dos falantes. Somada a isso, há o fato de que o acesso à educação na primeira metade do século XX estava restrito a uma minoria oriunda das famílias mais estabilizadas, economicamente. Como consequência, começa a configurar-se uma realidade polarizada em torno da língua: o português popular falado pela grande maioria da população que se encontrava longe dos bancos escolares, com uma cultura rural bastante arraigada; e o português culto falado por uma parcela minoritária que frequentava a escola, tendo, por conta disso, acesso à norma padrão, reproduzindo-a, parcialmente, na fala cotidiana. O *status* social do falante, inserido na cultura letrada, se encarregava de dar à norma culta uma feição que lhe era peculiar, distinguindo-a da norma popular e, até mesmo, da própria norma padrão, em muitos aspectos funcionais.

O quadro atual de acesso à escolarização, ao letramento mudou radicalmente num intervalo de pouco mais de 50 anos, porém, muitas lacunas precisam ainda ser preenchidas para minorar a situação de exclusão em que vivem milhares de pessoas que não tiveram a oportunidade de estudar. Basta lembrar que numa população de 306.866 habitantes, segundo o censo demográfico de 2010, com estimativa para 340,199 em 2014, Vitória da Conquista ainda ressoa o flagelo do analfabetismo que se abate sobre 65.011 moradores.

3.1.2 A pesquisa de campo

Como pudemos ver até agora, a pesquisa de campo constitui-se numa atividade essencial para a realização de um estudo no modelo laboviano, já que o intento é analisar o vernáculo (o como se diz, não o que se diz) num ambiente que venha a diminuir a presença do agente externo (o observador/pesquisador). Tal situação denominada por Labov (1972) de paradoxo do observador foi enfrentada pelos pesquisadores com a realização das entrevistas em espaços familiares aos informantes.

Após o preenchimento de um questionário com questões gerais, cujo objetivo era verificar se o informante encaixava-se no perfil desejado, passou-se à realização das

entrevistas. Tal tarefa coube aos bolsistas dos Grupos de Pesquisa em Linguística Histórica e Sócio-funcionalismo, ambos cadastrados no CNPQ, alocados na Universidade Estadual do Sudoeste da Bahia, *Campus* de Vitória da Conquista.

Assim, a pesquisa de campo foi realizada em Vitória da Conquista, zona urbana, considerando falantes de duas normas: a popular e a culta. Para nossa pesquisa, foram selecionados 12 (doze) falantes do Português culto cujos dados por nós analisados foram comparados com os dados analisados por Guimarães (2014), relativos ao português popular.

3.1.3 Constituição do *corpus*

Segundo Guimarães (2014), análises centradas na Sociolinguística exigem que o pesquisador faça:

- a) observação de uma comunidade de fala;
- b) relação de hipóteses a fim de executar o trabalho de campo;
- c) seleção dos informantes;
- d) constituição do *corpus*;
- e) análise dos dados coletados;
- f) interpretação à luz da Teoria da Variação.

No item observação da comunidade de fala, consideramos a convivência de duas normas (a popular e a culta) seguindo a tese de bipolarização proposta por Lucchesi (2000). Dessa forma, houve a necessidade de organizar o *corpus* considerando membros equânimes das normas consideradas. Logo, o reconhecimento de tais normas levou-nos a postular que a concordância de número no SN se realizaria de forma quantitativamente diferenciada e que tal diferenciação deveu-se à constituição sócio-histórica das duas normas, bem como à ação da cultura escolar. Em ambos os casos, haveria, também, fatores estruturais a condicionarem a marcação progressiva dos elementos constituintes do SN.

Além de considerar a constituição sócio-histórica das duas normas e elencarmos as hipóteses decorrentes de sua atuação no fenômeno da variação de número no SN, resolvemos agrupar os informantes segundo escolaridade, faixa etária e sexo. Esses condicionantes do comportamento social, sem dúvida, a nosso juízo, configuram-se, de mesma sorte, condicionantes do comportamento linguístico.

Devemos, agora, portanto, adentrar em questões relacionadas ao perfil dos informantes, etapa crucial na constituição do *corpus*.

3.1.4 Perfil dos informantes

Conforme já discutido na subseção 3.1.1, a constituição da comunidade de fala de Vitória da Conquista – BA está relacionada, inicialmente, à colonização portuguesa, resultado de um movimento expansionista com consequente dizimação dos primeiros habitantes da terra. Nesse sentido, podemos visualizar duas vertentes constitutivas da sociedade brasileira. A primeira está ligada ao aludido movimento expansionista o qual conduziu para o interior da Colônia membros da sociedade portuguesa. A segunda estaria ligada ao princípio fundador de nossa colonização: a servidão. Tal servidão seria materializada na domesticação do aborígine e na introdução de elementos adventícios, isto é, a importação de “peças³⁴” de africanos. Desse modo teríamos, em solo conquistense, a seguinte equação: brancos europeus e seus descendentes, incluindo, paulatinamente os pardos, e negros e índios domesticados, cujos descendentes constituiriam a base de nossa pirâmide social. Ambos os grupos estariam diferenciados pela educação formal e pelo acesso às profissões prestigiadas, bem como os postos de comando.

Assim posto, consideramos como pertencentes à norma popular os informantes que possuíam, no máximo, 5 anos de escolarização, de maneira que estão inseridos na “*parcela da população em que o letramento tenha sido precário ou não tenha ocorrido*” (SILVA, 2005, p. 148). Em relação aos informantes da norma culta, incluímos apenas aqueles que possuíam mais de 11 (onze) anos de formação escolar, considerando a educação profissional técnica e o nível superior.

A seguir, apresentamos um quadro contendo uma síntese comparativa dos informantes do português culto.

Quadro 3 – Informantes do Português Culto em Vitória da Conquista³⁵

Informante	Sexo	Código	Idade	Escolaridade
Faixa I				
F.S.L.B	M	PCVC01	19 anos	Nível superior
L.C.S	F	PCVC02	20 anos	Nível superior
C.B.S	F	PCVC03	21 anos	Nível superior
J.L.S	M	PCVC04	21 anos	Curso profissionalizante

³⁴ Tal nomenclatura refere-se aos negros assujeitados trazidos de África e assim nomeados por documentos oficiais, conforme Silva (2003).

³⁵ Elaborado com base na proposta de “ficha do informante” construída pelos professores Dr. Jorge Augusto Alves da Silva e Dr^a Valéria Viana de Sousa, coordenadores do Grupo JANUS – Grupo de Pesquisa em Linguística Histórica e Sociofuncionalismo (UESB) para ser aplicada nas atividades de pesquisa de campo quando da constituição do *corpus* (PPVC).

P.G.L	F	PCVC05	23 anos	Nível superior
A.F.S.F	M	PCVC06	26 anos	Curso profissionalizante
Faixa II				
H.F.D.S	M	PCVC07	36 anos	Nível superior
A.S.A	F	PCVC08	39 anos	Nível superior
E.S.F	F	PCVC09	47 anos	Curso profissionalizante
K.A.C	F	PCVC10	50 anos	Curso profissionalizante
D.A.O	M	PCVC11	51 anos	Curso profissionalizante
J.V.B	M	PCVC12	54 anos	Curso profissionalizante

3.1.5 As entrevistas

Seguindo o percurso da pesquisa de campo, após a escolha da comunidade, considerando sua relevância sócio-histórica, acreditando numa bipolarização linguística, decorrente de uma bipolarização social, passamos à escolha dos informantes, considerando o perfil que se encaixasse nos critérios sociais elencados, a saber: local de origem, idade, sexo e escolarização. Tais critérios ficaram explícitos no “questionário social” e se tornaram instrumentos balizados da escolha dos informantes. Após tal tarefa, passamos às fases das entrevistas, que ora transcrevemos, parafraseando, o que consta em Guimarães (2014):

1º Passo: Contato direto e prévio com os informantes a fim de verificar se correspondiam ao perfil desejado.

2º Passo: Aplicação do questionário: nome completo, idade, sexo, filiação, profissão, nível de escolaridade, viagens, penetração dos meios de comunicação no ambiente doméstico.

3º Passo: Realização da entrevista: zona urbana, local próximo à moradia ou na moradia do informante, entrevista de 50 minutos (em média), documentador entrevista informante. Para tal fim, foram utilizados minigravadores digitais Coby CXR 190 e Olympus Vn – 500.

Todas as entrevistas foram repassadas para arquivos próprios e, hoje, fazem parte do acervo do Grupo de Pesquisa em Linguística Histórica e Sociofuncionalismo e integram também o acervo do Projeto *Vertentes*, coordenado por Dante Lucchesi – UFBA.

Guimarães (2014) observa que:

As entrevistas foram realizadas com a finalidade de que o informante se comportasse o mais natural possível, no sentido de serem conduzidas em um tom com característica bastante informal, uma conversa. Mesmo portando um roteiro visível aos falantes, conduzimos a entrevista num formato de conversa com certo grau de informalidade. O falante foi estimulado a falar de assunto/temática/algo de que tinha conhecimento, ou que o houvesse

marcado emocionalmente. Com isso, esperávamos que as presenças do gravador e do documentador fossem, minimamente, deixadas de lado, possibilitando maior uso do vernáculo pelo falante. Tal metodologia procura atender ao requisito laboviano de representar o vernáculo da comunidade de fala.

Tais práticas elencadas por Guimarães (2014) visam à tentativa de diminuir o paradoxo do observador. A mesma pesquisadora elenca os assuntos mais profícuos para a conversação-entrevista, variando os interesses conforme a faixa etária e o sexo. No geral, pudemos verificar que lembranças da infância, casos locais (crimes, roubos) e práticas religiosas foram os temas mais recorrentes para ambos os sexos e para as duas faixas etárias³⁶.

3.1.6 Processo de coleta e armazenamento de dados

Após a realização das entrevistas, passamos à execução das etapas de audição e transcrição. A audição das fitas e a transcrição foram feitas por graduandos do curso de Letras da UESB. O trabalho de revisão das transcrições (cotejo entre o áudio e a transcrição) foi feito por bolsistas e por membros da iniciação científica, todos vinculados aos Grupos de Pesquisas em Linguística Histórica e Sócio-funcionalismo, ambos liderados pelos professores doutores Jorge Augusto Alves da Silva e Valéria Viana Sousa.

Para efetivação da transcrição, foi usada a “Proposta de Chave” elaborada pelo professor Doutor Dante Lucchesi³⁷ (UFBA).

Dada a especificidade da nossa pesquisa, demos atenção especial à presença de marcas formais de regras de concordância nominal de número no SN.

3.1.6.1 Critérios da transcrição

Com base no estudo realizado por Silva (2005), optamos pela transcrição grafemática; no entanto, por tratarmos de um estudo morfossintático, algumas características gerais da língua portuguesa do Brasil, mesmo que não estivessem em acordo com a norma gramatical vigente, não foram destacadas.

Na transcrição das gravações, foram desprezados alguns fatos linguísticos como:

³⁶ No conjunto do *corpus*, foram consideradas duas faixas etárias: Faixa I (20 a 35 anos); Faixa II (36 a 50 anos).

³⁷ Líder do grupo VERTENTES/UFBA, a quem muito devemos, tamanho é o apoio concedido à realização dos trabalhos pertinentes ao grupo de pesquisa Janus/UESB, bem como à concessão de instrumentalização ao grupo de pesquisa, mencionado, como um todo, de modo especial, quando da oferta de curso junto a outros membros do grupo VERTENTES no período da realização do I Seminário de Variação e Mudança Linguística no Sudoeste – BA, novembro de 2013/UESB.

- a) A elevação das vogais médias em posição átona final;
- b) A ditongação antes de consoante constrictiva implosiva;
- c) A palatalização do /t/ e /d/, antes de vogal palatal;
- d) A epêntese da vogal alta que desfaz o travamento do grupo consonântico;
- e) A vocalização da consoante lateral pós-vocálica.

3.1.7 Levantamento de dados

No levantamento dos dados, procuramos colher exaustivamente as ocorrências, observando como os elementos flexionáveis da concordância de número no sintagma nominal se realizavam. A fim de mantermos o rigor científico, mesmo assumindo a competência das transcrições, tivemos que recorrer, em alguns casos, à audição das gravações para sanarmos algumas dúvidas.

Para a análise da recorrência da *concordância nominal de número* quando do uso de marcas formais nos elementos do SN, na comunidade de fala de Vitória da Conquista, após processo da transcrição e levantamento de ocorrências já informadas, destacamos todos os sintagmas nominais pluralizáveis constituídos de pelo menos dois elementos em condição de solidariedade. Em seguida, procedemos à codificação dos dados com o auxílio da “Chave de Análise Mórfica” proposta pelo professor Dante Lucchesi.

Como ferramenta de quantificação, usamos a versão GoldVarDb 3b.

3.2 A VARIÁVEL DEPENDENTE E AS VARIÁVEIS INDEPENDENTES

No Português do Brasil, a concordância nominal de número encontra-se em variação. Isso significa que podemos verificar a marcação em um ou em alguns dos elementos do SN ou até mesmo a marcação de número em todos os elementos existentes, embora tais marcas sejam até redundantes (como no caso do plural metafônico). Nesse sentido, a alternância de uso da *concordância de número no sintagma nominal*, no português brasileiro, indica que o falante tende a lidar quase que cotidianamente com ocorrências sintagmáticas em que os constituintes ora são redundantemente marcados na sua forma plural, ora recebem a indicação de pluralidade em alguns ou apenas um desses elementos.

Nos diversos estudos sobre a variação na concordância nominal no SN, já realizados no Brasil e apresentados ou referenciados ao longo desse trabalho, o que podemos notar é a

enorme relevância em torno de algumas variáveis linguísticas e extralinguísticas condicionadoras do fenômeno. Nesse momento, discutiremos aquelas controladas em muitas das pesquisas por nós aduzidas a fim de observarmos a atuação de tais variáveis nos dados de fala por nós investigados. Partindo sempre de uma perspectiva atomística, nossa abordagem recairá sobre as variáveis estruturais: posição linear do constituinte, posição do constituinte com referência ao núcleo do SN, classe gramatical do constituinte e saliência fônica; assim como as variáveis sociais: faixa etária, sexo e escolaridade.

3.2.1 Variáveis linguísticas ou estruturais

Em nossa pesquisa, consideramos as mesmas variáveis linguísticas ou estruturais controladas por Guimarães (2014) na análise dos dados do português popular de Vitória da Conquista, tendo em vista que os resultados por ela encontrados serão comparados com os nossos, relacionados ao português culto. Abaixo, apresentamos cada uma dessas variáveis.

3.2.1.1 Posição linear do constituinte

Começamos pela variável posição linear, partindo da hipótese de que os constituintes que ocupam a primeira posição são sempre mais marcados formalmente do que os demais elementos do sintagma nominal. Inserimo-nos, portanto, no contexto dos estudos, como o de Martins (2013), segundo o qual “quanto mais à direita o elemento linguístico estiver no SN, sendo núcleo ou não, há um desfavorecimento no uso da aplicação da regra” (MARTINS, 2013, p. 143).

Os estudos de Scherre (1988), Fernandes (1996), H. Carvalho (1997) e R. Carvalho (1997), quando analisaram essa variável isoladamente, confirmaram, igualmente, o favorecimento da aplicação da regra no constituinte que ocupa a primeira posição. Portanto, a caracterização da variável posição linear se dá, conforme Schneider (2012), “em função do local que o elemento analisado ocupa no sintagma nominal, no sentido estritamente linear” (SCHNEIDER, 2012, p.125).

Os dados por nós codificados para essa variável foram organizados de acordo com a exemplificação do quadro 3³⁸:

³⁸ Por questão metodológica, apenas indicaremos as fontes para as figuras e as tabelas que não forem de nossa autoria, subentendendo-se, portanto, que as demais são de nossa lavra.

Quadro 4: Grupo de fatores da variável *posição linear do constituinte*

Fatores	Exemplos
Itens na primeira posição do SN	<i>As criança</i> Ø tava brincando...
Itens na segunda posição do SN	Umas <i>dança</i> Ø...
Itens na terceira posição do SN	Umas praia <i>limpa</i> Ø...
Itens na quarta posição em diante	Os meus dois <i>tio</i> Ø...

Nesse sentido, postulamos quatro fatores que foram observados, codificados e quantificados, segundo Chave de Codificação. Ademais, a escolha de tais fatores segue a proposta feita por Schneider (2012).

3.2.1.2 Natureza³⁹ do constituinte e sua posição em relação ao núcleo do SN

Em sua pesquisa, Scherre (1988) demonstrou, com muita propriedade, que os constituintes sintagmáticos que antepõem o núcleo são sempre os mais marcados. H. Carvalho (1997) e Lopes (2001) também chegaram à mesma conclusão. Brandão (2011) considerou tal variável, chegando à constatação de que a maior incidência de plural no núcleo ocorre quando o elemento nuclear ocupa a primeira posição. Nossa hipótese é a de que os determinantes em primeira posição à esquerda do núcleo são os mais marcados formalmente.

Para essa variável, estabeleceu-se a seguinte organização, conforme quadro 5:

Quadro 5: Grupo de fatores da variável *posição do constituinte com referência ao núcleo*

Fatores	Exemplos
Determinante na primeira posição imediatamente à esquerda do núcleo.	MEUS filhos tudo adoeceu; MEUS filho
Determinante em primeira posição, mas não adjacente (não imediatamente) ao núcleo.	MEUS primeiros filhos; eu gosto dOS meus filho
Determinante em segunda posição imediatamente à esquerda do núcleo.	Todos MEUS filhos; eu gosto dos MEUS filho.
Determinante em segunda posição não adjacente ao núcleo.	Todos MEUS melhores amigos; Todos MEUS melhor amigos
Determinante em terceira posição à esquerda do núcleo.	Todos meus MELHORES amigos; Todos meus MELHOR amigo.
Núcleo em primeira posição.	Vi MISSAS belíssimas; Comprei bananas madura.

³⁹ Em nosso estudo, entendemos como natureza a possibilidade de o elemento estrutural funcionar como núcleo, determinante e quantificador, tal qual doutrina defendida por Lopes (2001).

Núcleo em segunda posição .	Meus FILHOS; Meus PARENTE.
Núcleo em terceira posição	E os meus FILHOS casaram; Os oto IRMÃO.
Modificador (constituente) imediatamente à direita do núcleo)	Eu gosto de festas RELIGIOSAS; Meus irmãos maió.
Modificador (constituente) à direita do núcleo mas não adjacente a ele (não imediatamente a ele)	(Meus filhos mais VELHOS; Meus irmãos mais CHEGADO

No âmbito desse estudo, realizamos uma abordagem atomística e em decorrência disso, os sintagmas nominais são analisados em conformidade com sua natureza e sua posição. Nesse sentido, alguns amálgamas podem ser efeitos para dar conta de ajustes necessários à interpretação da variável.

3.2.1.3 Classe gramatical do constituinte

A variável *classe gramatical* tem a ver com a natureza morfológica de cada constituinte e sua relevância conjugada ao grau de iconicidade e processamento cognitivo. Schneider (2012), partindo de uma visão crítica ao trabalho de Scherre (1988), destaca que esse tem sido um aporte funcionalista usado por aquela pesquisadora, acreditando na importância semântica na aplicação da marcação de número no SN. Assim, o resultado encontrado por Scherre (1988), em seus estudos, revelou que a ocorrência envolvendo os substantivos no sintagma nominal pode se dar na primeira, na segunda e na terceira posição. A autora observou, com isso, que os substantivos na primeira posição, na condição de núcleo, favorecem muito mais a marcação de plural; e que os substantivos na terceira posição são até mais marcados do que os da segunda posição. Já os adjetivos, embora ocorram nas três posições, os da segunda tendem a receber mais marcas formais do que os da terceira, de acordo com Scherre. A pesquisadora destaca, ainda, a atuação dos pronomes indefinidos e dos quantificadores que transitam, igualmente, nas três posições do sintagma ao passo que os possessivos ocorrem na primeira e na segunda posição. Nossa hipótese, no entanto, é a de que os artigos definidos e indefinidos assim como os pronomes demonstrativos, por atuarem como determinantes no SN, tendem a ser muito mais marcados formalmente do que as demais classes gramaticais. O Quadro 6 abaixo exemplifica melhor a formação de tal variável.

Quadro 6: Grupo de fatores da variável *classe gramatical do constituinte*

Fatores	Exemplos
Artigo definido	<i>Os meus irmão...</i>
Artigo indefinido	<i>Umas manhã...</i>
Pronome demonstrativo	<i>Essas coisa...</i>
Pronome possessivo	Com <i>meus</i> amigo...
Pronome indefinido	<i>Alguns</i> preços justo...
Quantificador	<i>Todos</i> os dias...
Substantivo	Mais de três <i>dia</i> ...
Adjetivo	Minhas amizades <i>verdadeira</i> ...

Acreditamos, contudo, que essa variável esteja concorrendo com a natureza e posição do constituinte em relação ao núcleo do SN.

3.2.1.4 Saliência fônica

Esse é um princípio segundo o qual “as formas mais salientes, e por isso mais perceptíveis, são mais prováveis de serem marcadas do que as menos salientes” (SCHERRE, 1988, p. 64). O que está em jogo, na verdade, é a diferenciação de material fônico na oposição singular/plural, como é o caso de *novo/novos*, *hotel/hotéis*, por exemplo.

Filiando-nos, do mesmo modo, aos estudos de Fernandes (1996), H. Carvalho (1997), R. Carvalho (1997), Lopes (2001), Silva (2011), Martins (2013), entre outros, esperamos, em nossa pesquisa, que os itens lexicais com formação de plural irregular, em sua grande maioria, favoreçam mais a indicação de pluralidade por serem mais perceptíveis na relação singular/plural.

Organizamos a variável em questão, considerando os fatores presentes no quadro 7.

Quadro 7 – Grupo de fatores da saliência fônica

Plural duplo ou metafônico (ex.: ovo/ovos)
Alteração do ditongo [aw/ay] (ex.: animal/animais)
Alteração de um ditongo, inclusive da vogal (ex.: leão/leões)
Alteração de um ditongo, mantendo a vogal (ex.: pão/pães);
Acréscimo do –s (ex.: tom/tons; irmã/irmãs; algum/alguns; irmão/irmãos)
Acréscimo do –s, formando sílaba, reconstituindo elemento (ex.: cantor/cantores)
Acréscimo do –s/ -z, formando sílaba sem reconstituição (ex.: freguês/ fregueses; vez/vezes)

Acréscimo de -s (ex.: casa/casas)
Acréscimo de -s, formando ditongo (armazém/armazéns)

A variável saliência fônica tem sido selecionada em quase todos os trabalhos sobre a concordância nominal. Tal variável conjuga, além do material mórfico envolvido, a tonicidade como elemento coadjuvante na aplicação da regra. Assim, no conjunto dos fatores elencados, podemos ver que foram consideradas as possibilidades de existência de palavras acentuadas e não acentuadas, bem como o envolvimento de mudança ou acréscimo de sílaba, alterando o material fônico, tal como ocorre em *freguês* > *fregueses*.

No âmbito de nosso estudo, fizemos uma interpretação daquilo que tem sido usado pela literatura já clássica no assunto, reinterpretando à luz dos estudos realizados por Guimarães (2014).

3.2.1 Variáveis extralinguísticas ou sociais

No âmbito desse estudo, consideramos três variáveis extralinguísticas ou sociais, a saber: sexo, escolaridade e faixa etária.

3.2.1.1 Sexo

Na maioria das pesquisas sociolinguísticas que lidam com a concordância nominal de número, o que se tem observado é que as mulheres fazem mais uso das formas de prestígio do que os homens. O que pode ser atestado, por exemplo, nos trabalhos de Scherre (1988), Fernandes (1996), Martins (2013), entre outros. Por outro lado, alguns outros estudos realizados por R. Carvalho (1997) e Silva (2011) indicam que são os homens quem aproximam mais das formas de prestígio.

Nossa hipótese é a de que as mulheres tendem a fazer mais concordância nominal de número do que os homens, considerando que elas são muito mais atentas às convenções sociais, incluindo, no caso, os padrões de uso da língua. A variável foi organizada, conforme os fatores: (F) feminino e (M) masculino.

3.2.1.2 Escolaridade

As diversas pesquisas já citadas por nós nesse trabalho são unânimes em reconhecer que quanto mais escolarizado o indivíduo, mas ele tende a se aproximar das formas de prestígio. Inclusive, partimos da hipótese de que a aplicação da regra de concordância nominal é maior entre os falantes com grau de escolaridade em nível superior uma vez que eles estão muito mais expostos a atividades culturais diversificadas nas quais predominam o uso das formas linguísticas mais prestigiadas.

Organizamos essa variável, em nosso estudo, considerando os seguintes níveis de escolarização:

- ✓ Mais de onze anos de escolaridade: formação em curso profissionalizante.
- ✓ Mais de onze anos de escolaridade: formação em nível superior.

Nesse sentido, acreditamos que os anos de estudo podem alterar o vernáculo de falantes, devido ao contato com formas prestigiadas.

3.2.1.3 Faixa etária

O que se tem observado em muitas análises sociolinguísticas realizadas em estudos que lidam com a variação na concordância nominal é que, em se tratando da variável faixa etária, temos, de um lado, um padrão curvilíneo em que a indicação é justamente de variação estável, conforme se pode notar em Scherre (1988), por exemplo; de outro lado, há uma indicação de mudança em curso rumo a um sistema com ou sem presença de marcas formais de plural, sugerindo, com isso, um padrão linear. As pesquisas de Fernandes (1996), Lopes (2001), H. Carvalho (1997), Brandão (2011), entre outras, se inserem em tal padrão.

Em nossa pesquisa, organizamos essa variável considerando duas faixas: F I (20 a 35 anos) e F II (36 a 50 anos). Esperamos, com isso, que entre os informantes do português culto, os adultos façam mais uso das marcas de plural no SN do que os jovens.

Descrevemos, nesse capítulo, os procedimentos metodológicos de nosso estudo que abrangeu tanto o conhecimento da formação sócio-histórica da comunidade de fala por nós pesquisada, quanto o passo-a-passo da pesquisa. Além disso, apresentamos as variáveis linguísticas e extralinguísticas, tecendo algumas considerações sobre cada uma delas e levantando, ao mesmo tempo, as nossas hipóteses. No capítulo a seguir, iremos observar se elas foram ou não atestadas com base nos nossos resultados probabilísticos.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Apresentamos, nesse capítulo, a discussão referente aos resultados da análise estatística da variável dependente: *concordância nominal de número no SN*. Os dados da análise foram extraídos do *corpus* do Português Culto de Vitória da Conquista (PCVC) e cotejados com o *corpus* do Português Popular de Vitória da Conquista (PPVC), utilizando para tal comparação a pesquisa realizada por Guimarães (2014). Os *corpora* por nós utilizados foram construídos pelos Grupos de Pesquisa em Linguística Histórica e em Sociofuncionalismo da Universidade Estadual do Sudoeste da Bahia – UESB.

Cabe-nos esclarecer que, na presente análise, investigamos o fenômeno da concordância nominal de número no SN, numa **abordagem atomística**, considerando doze informantes, os quais divididos segundo critérios sociais em faixa etária e gênero/sexo apresentam em comum o fato de possuírem mais de onze anos de formação escolar. Nesse sentido, o vernáculo de tais informantes foi comparado, usando critérios idênticos aos empregados por Guimarães (2014), exceto no que concerne à escolaridade, já que aquela pesquisadora avaliou seus informantes, considerando os “sem escolarização” e os “com escolarização precária”. Em suma, os dados ora apresentados foram obtidos a partir das seguintes realidades:

- a) Análise de amostras de fala de 12 informantes do *corpus* PCVC, agrupados segundo critérios sociais (faixa etária, gênero/sexo, nível de escolaridade);
- b) Cotejo dos resultados obtidos com a análise empreendida, tendo como elemento de comparação o estudo realizado por Guimarães (2014).

O comportamento linguístico dos falantes de tais normas⁴⁰ foi mensurado, nesse estudo, a partir da ocorrência de indicação de pluralidade em cada constituinte do SN, o que faz o presente estudo encaixar-se numa perspectiva analítica dita mórfica ou **atomística**.

Pretendemos, igualmente, atestar a hipótese de que a **variação na concordância nominal** entre os elementos flexionáveis do SN é maior entre os falantes da norma popular uma vez que os falantes da norma culta estão ou estiveram muito mais expostos à cultura letrada em função de já terem cursado mais de onze anos de escolarização em níveis de formação técnica profissional e superior, o que os fizeram se aproximar das formas mais prestigiadas, linguisticamente. Nesse sentido, ainda, nosso estudo corrobora a hipótese de

⁴⁰ No âmbito de nosso estudo, as expressões “norma culta” e “norma popular” *assim* são tomadas como equivalentes a Português **Culto** e Português **Popular** tal qual advoga Lucchesi (1994) a partir de releitura de Coseriu.

realidade bipolarizada (LUCCHESI, 2000) e que tal bipolarização poderia facilmente ser explicada pela constituição histórica do Brasil e de suas variedades de português.

Em síntese, a hipótese norteadora de nosso estudo pauta-se na convicção de que os utentes da norma intitulada Português Culto de Vitória da Conquista apresentem, em termos probabilísticos, maiores ocorrências de **concordância nominal de número no SN** graças ao contato com aparatos sociais advindos de uma cultura letrada, escolar em sua essência, cuja sócio-história reflete os valores de uma parcela significativa da comunidade conquistense. Para tanto foram controlados grupos de fatores linguísticos e extralinguísticos, submetidos à análise estatística probabilística, usando-se a ferramenta GoldVarb 3.0b3. Para tanto, foram considerados significativos os seguintes grupos de fatores, com Input .988 e significância de .004:

- a) Posição linear do constituinte;
- b) Posição do constituinte com referência ao núcleo do SN;
- c) Saliência fônica;
- d) Gênero/Sexo;
- e) Faixa etária.
- f) Escolaridade

Em seguida, passaremos a discutir os grupos de fatores, comparando-os aos de Guimarães (2014).

4.1 A VARIAÇÃO NA CONCORDÂNCIA NOMINAL DE NÚMERO NO SINTAGMA NOMINAL NO PORTUGUÊS POPULAR E CULTO DE VITÓRIA DA CONQUISTA

No âmbito dessa pesquisa, a análise quantitativa dos dados envolveu um total de 2.205 ocorrências extraídas de amostras de fala de doze informantes do PCVC. Como se trata de uma análise atomística, foram considerados os SN com todas as marcas formais de plural, SN com algumas marcas ou apenas uma marca de plural e SN contendo um numeral como primeiro elemento. Por outro lado, desconsideramos as ocorrências em contextos linguísticos em que não se podia verificar a presença do plural no elemento seguinte⁴¹, bem como aqueles SN cujos núcleos são ditos invariáveis⁴². A **Tabela 4** expressa a realidade mensurada por meio da ferramenta GoldVarb 3.0b3.

⁴¹ Eis um exemplo: “as meninas são amigas”.

⁴² Eis um exemplo: “ele pegou dois ônibus”.

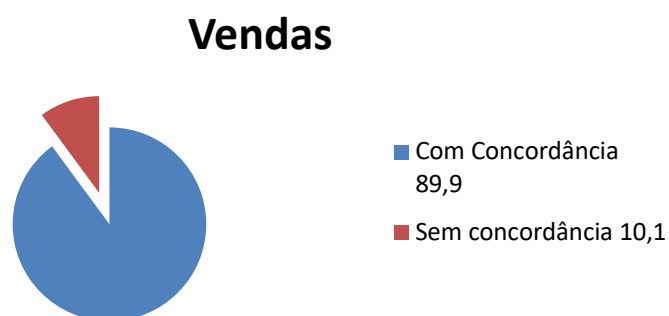
Tabela 4 – Variável dependente: Concordância nominal de número

Concordância Nominal	Português culto	%
Com marcas de Plural	1.982 / 2.205	89,9%
Sem marcas de Plural	223 / 2.205	10,1%
Total de Ocorrências	2.205	

Os dados demonstram, portanto, que a concordância de número no SN das amostras de fala de utentes de Vitória da Conquista – BA assemelha-se ao percentual encontrado por Lopes (2001) na capital do Estado (Salvador). Cumpre lembrarmos que Lopes faz diferenciação entre nível de escolaridade, considerando falantes de 11 anos e de no mínimo 15 anos. Os dados obtidos por Lopes (2001, p.163) foram:

- a) Escolaridade média (11 anos) 82%;
- b) Escolaridade superior (mínimo de 15 anos) 96%.

Eis o **Gráfico 1** elaborado a partir dos dados obtidos:

Gráfico 1: Porcentagem de concordância no SN - PCVC

No cotejo com os dados obtidos por Guimarães (2014), das 2.979 ocorrências que foram extraídas da fala de doze informantes do PPVC, os resultados são esses que ora apresentamos, em comparação com os dados por nós obtidos.

Tabela 5 – Variável dependente: Concordância nominal de número no PPVC e no PCVC

Concordância Nominal	PPVC	%	Português culto	%
Com marcas de Plural	1.708 / 2.979	57,3%	1.982 / 2.205	89,9%
Sem marcas de Plural	1.271 / 2.979	42,7%	223 / 2.205	10,1%
Total de Ocorrências	2.979		2.205	

Comparando os resultados obtidos com a variável dependente: concordância nominal de número, em ambas vertentes do português de Vitória da Conquista, conforme **Tabela 5**, já se pode confirmar, preliminarmente, a nossa hipótese, segundo a qual os falantes da norma

culta fazem mais concordância nominal de número do que os falantes da norma popular, considerando que os falantes da norma culta foram mais expostos à educação formal, resultado não só do acesso ao letramento, mas de uma perspectiva de mundo em que a escola passa a ser uma fonte de ascensão social. Embora essa constatação possa soar como um truísmo em função, sobretudo, dos diferentes níveis de escolaridade dos informantes, não pode servir de pretexto para se ignorar a atuação das variáveis linguísticas e extralinguísticas sobre o fenômeno por nós investigado. Além do mais, é preciso considerar que as normas espelham realidades de vida em que valores são postos como lastros de uma vida desejável.

A despeito da diferença entre as duas normas, torna-se oportuno trazer a discussão feita por Guimarães (2014) a qual considera fundamental observar o aparato urbanizador na expansão territorial e expansão simbólica (considerando-se a educação como símbolo de prestígio). Seguindo esse raciocínio, espelhar-se na linguagem da escola torna-se um bem de consumo para as classes sociais (quer antigas quer novas) que ocupam os espaços privilegiados (territoriais, simbólicos e econômicos) em Vitória da Conquista. O *locus urbanus* seria, portanto, o espaço de distribuição simbólica em que a escola entraria como elemento balizador e sua linguagem referendaria o *locus urbanus* assumido.

Nesse sentido, podemos comparar os nossos dados e os obtidos por Guimarães (2014) a outros estudos realizados em condições sócio-históricas assemelhadas, como no caso do trabalho realizado por Lília Soares Miranda Santos na cidade de São Leopoldo –MG. Ao pesquisar as normas popular e culta, Santos (2010) encontrou realidade semelhante à vista em Vitória da Conquista no que concerne à norma culta, já que constatou 75% de concordância no SN, mas apenas 24% na amostra a qual denominou de Fundamental. A investigação sociolinguística nos grandes centros urbanos, como nas capitais, a nosso juízo, apresenta um estrato populacional significativo, mas insuficiente para demonstrar a nossa pluralidade linguística. É nesse sentido que estudos realizados por Silva (2005), Santos (2010), Martins (2013) e Guimarães (2014) vêm preencher lacunas em relação às novas realidades que surgem com o crescimento populacional e econômico dos novos polos urbanos.

Em seguida, passamos a apresentar e a discutir os grupos de fatores selecionados pelo GoldVarb 3.0b3, observando a ordem de seleção dos grupos considerados relevantes. Nesse sentido, julgamos necessário proceder à análise daquelas variáveis que foram selecionadas como relevantes, do ponto de vista estatístico, trazendo os dados do PCVC, comparando-os aos dados do PPVC presentes no trabalho realizado por Guimarães (2014). Em nosso recorte, preferimos destacar também a realidade exposta pelos dados de Santos

(2010), considerando a semelhança do processo de urbanização que há entre Vitória da Conquista (interior da Bahia) e São Leopoldo (interior de Minas Gerais).

A seguir, são apresentados e discutidos os resultados referentes às variáveis:

- g) Posição linear do constituinte;
- h) Posição do constituinte com referência ao núcleo do SN;
- i) Saliência fônica;
- j) Gênero/Sexo;
- k) Faixa etária.

Reiteramos que nosso estudo tem como lastro a comparação com os resultados obtidos por Guimarães (2014), a qual realizou estudo apresentando evidências sobre a configuração sócio-histórica do português popular de Vitória da Conquista. Tal estudo integra as investigações iniciadas pelo Grupo Janus/CNPq.

Seguindo a linha de raciocínio de Guimarães (2014), devemos também interpretar os dados não selecionados a fim de retirar deles algumas lições que vão além da quantificação. Assim, embora a variável classe gramatical do constituinte não tenha sido selecionada pelo programa, apresentaremos os resultados percentuais e a análise interpretativa para essa variável.

4.1.1 Posição linear do constituinte

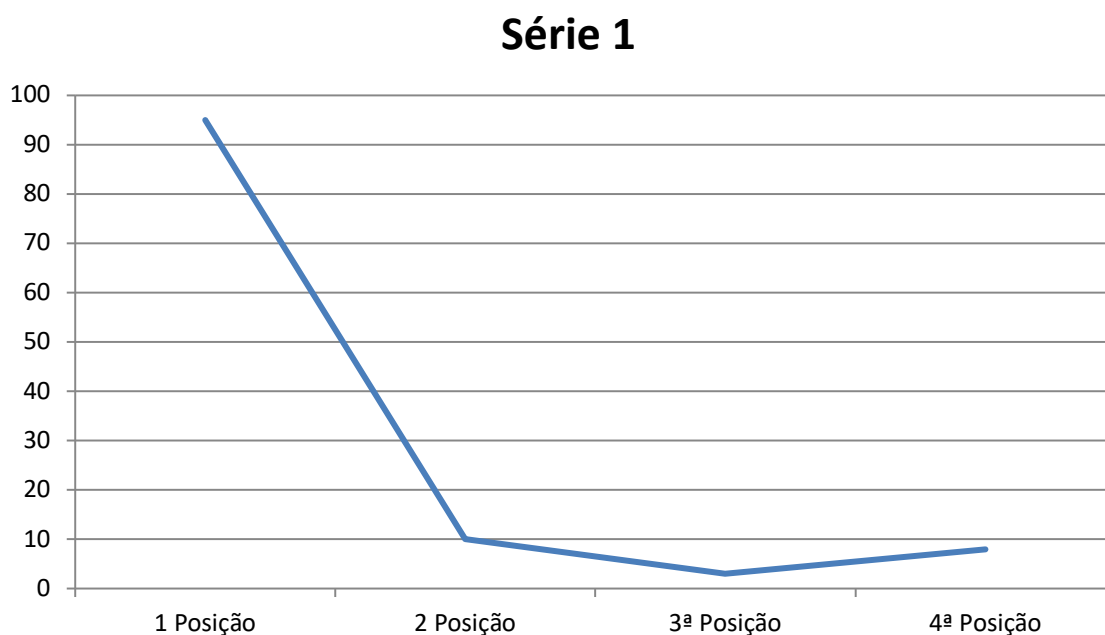
Para essa variável, nossa hipótese é a de que a primeira posição seja mais favorável à presença de marcas de plural, considerando o estabelecimento da oposição singular vs plural. Para tanto, consideramos, seguindo o estudo de Guimarães (2014), quatro fatores, como tivemos oportunidade de discutir anteriormente.

Apresentamos, mormente, os dados por nós obtidos da análise do PCVC a fim de comparar os resultados aos do PPVC.

Tabela 6 - Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável *posição linear do constituinte*.

Posição linear do constituinte	Português culto		P.R.
Primeira posição	993/995	99,8%	0,95
Segunda posição	853/1.046	81,5%	0,10
Terceira posição	131/158	82,9%	0,03
Quarta posição em diante	5/6	83,3%	0,08

Os resultados não deixam dúvidas de que a primeira posição tende a favorecer a retenção da marca de plural, o que nos sugere que tal fator, quando analisado isoladamente, conforme fizemos, acaba atestando o caráter funcional da língua, segundo o qual, para evitar redundância, há uma tendência para a informação relevante ser retida no primeiro constituinte sintagmático. A fim de melhorar a visualização, apresentamos o **Gráfico 2**.



Guimarães (2014), cujos dados do PPVC são aqui apresentados para compararmos com os dados do PCVC, lembra que as explicações e justificativas para a marcação do primeiro elemento do SN têm suscitado toda uma polêmica em torno das possíveis motivações para essa realidade. No bojo das discussões, estaria a origem do português popular do Brasil, envolvido num acirrado debate em que de um lado estaria a *deriva secular românica* e de outro a *transmissão linguística irregular*. Seguindo este último raciocínio, o contato maciço do português, sobretudo com as línguas africanas, teria desencadeado peculiaridades morfossintáticas, como a variação na concordância nominal de número, que refletiria tendências das línguas crioulas de indicarem o plural, por exemplo, apenas nas estruturas mais relevantes da língua. Por outro lado, segundo aquele postulado, as línguas românicas estariam caminhando para uma simplificação morfológica, pressionadas pelo desaparecimento do rígido sistema de casos do latim; assim, a marcação de apenas um elemento seria uma forma de resgatar, por um caráter sintético já previsto no Latim Vulgar, a informação necessária para estabelecer as oposições e marcá-las. Dito isto, tais posturas têm

sido levadas como divisores de água, atualmente em discussão que pretendem explicar a origem do Português do Brasil, conforme discutiu amplamente Araújo⁴³.

Empregando um aporte teórico Funcionalista, Scherre (1988) já veria a questão do favorecimento da marcação de plural na primeira posição como a atuação de um princípio funcional: as *Condições de Distintividade*. De acordo com tal princípio, há uma tendência para a informação semanticamente relevante ser retida na estrutura superficial da língua, o que faz com que informações redundantes sejam canceladas (SCHERRE, 1996, p. 92).

Nessa polêmica, nosso estudo se insere numa tentativa de discutir princípios sócio-históricos que serviriam para explicar não só a formação do espaço urbano, mas a formação do espaço urbano símbolo pela constituição da história linguística. Os núcleos populacionais do interior da Bahia surgem por meio de uma elite agrária de caráter expansionista, ocupando posteriormente o poder político local. O aparato urbanizador, nesse sentido, teria como mola mestra a formação de uma elite cultural, a qual teria a escola como combustível impulsionador.

Retornando à discussão linguística, os dados por nós apresentados para o PCVC estão em consonância com os estudos realizados por outros pesquisadores, tais como Lopes (2001), Andrade (2003), Santos (2010) e Martins (2013). Vejamos os dados encontrados pelas pesquisadoras:

- a) Andrade (2003), por exemplo, encontrou para a primeira posição peso relativo de 0.94. Ela analisou a fala de informantes de Helvécia (BA), uma comunidade rural de afrodescendente.
- b) Martins (2013) também encontrou peso relativo alto para a primeira posição, 0.79. A autora investigou a fala de habitantes da microrregião do alto Solimões (AM).
- c) A pesquisa de Lopes (2001) apresenta, igualmente, um alto índice de marcas de plural na primeira posição, com peso de 0.85. A pesquisadora analisou dados de fala do português popular e culto de Salvador.

Adentremos, agora, na comparação entre os dados do PCVC e do PPVC, observando a tabela produzida a partir do cotejo dos nossos dados com os de Guimarães (2014):

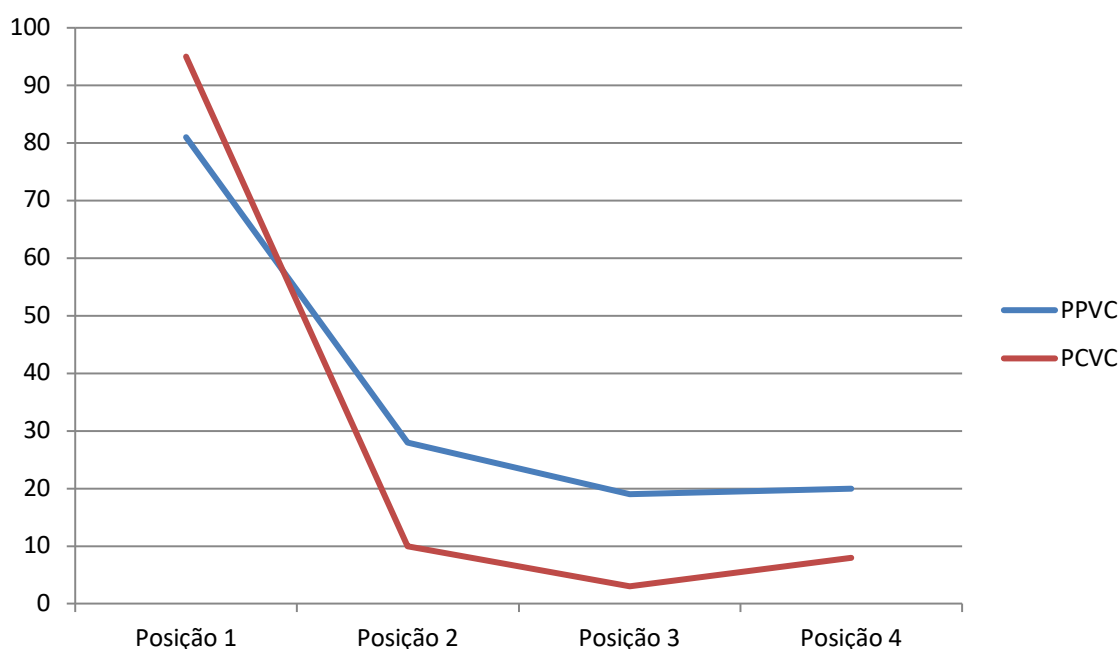
⁴³ ARAÚJO, S. S. de F. A concordância verbal no português falado em Feira de Santana – BA. Sociolinguística e sócio-história do português brasileiro. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Língua e Cultura, Universidade Federal da Bahia, Salvador, 2014.

Tabela 7: Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável *posição linear do constituinte*.

Posição linear do constituinte	PPVC	P.R.	Português culto	P.R.
Primeira posição	1.183/1.195 99%	0.81	993/995 99,8%	0.95
Segunda posição	455/1.522 29,9%	0.28	853/1.046 81,5%	0.10
Terceira posição	59/222 26,6%	0.19	131/158 82,9%	0.03
Quarta posição em diante	11/40 27,5%	0.20	5/6 83,3%	0.08

O gráfico 3 apresenta de forma sintética os pesos relativos referentes ao PPVC (GUIMARÃES, 2014) e ao PCVC.

Gráfico 3: Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável *posição linear do constituinte*.



Como podemos ver, é inegável o favorecimento da primeira posição quando se trata de concordância de número no SN. Julgamos, contudo, que a aplicação do princípio funcionalista de *Condições de Distintividade* não anularia o olhar sócio-histórico que lançamos sobre a questão, já que tal estratégia é factível em diversas situações linguísticas, como demonstrou Silva (2005). Nas demais posições ocupadas pelos constituintes por nós analisados, podemos observar que tanto no PPVC quanto no culto os pesos relativos apontam para o

desfavorecimento decrescente na indicação de pluralidade, na seguinte ordem: segunda posição 0.28 (PPVC) e 0.10 (PCVC), quarta posição 0.20 (PPVC) e 0.08 (PCVC), terceira posição 0.19 (PPVC) e 0.03 (PCVC).

Efeito semelhante ocorreu com os dados de Martins (2013) cujo desfavorecimento na aplicação da regra a partir da segunda posição é bastante notório, conforme os seguintes pesos: segunda posição 0.28, terceira posição 0.28 e quarta posição 0.20. Essa tendência decrescente também é encontrada na análise de Lopes (2001) cujos resultados probabilísticos revelam-nos pesos de 0.22 para a segunda posição, 0.21 para a terceira posição e 0.17 para quarta posição.

4.1.2 Posição do constituinte com referência ao núcleo

A hipótese por nós levantada é a de que os determinantes em primeira posição à esquerda do núcleo favorecem mais a marcação de plural.

Ao contrário de Guimarães (2014) que considerou cada um dos fatores da variável posição do constituinte com referência ao núcleo, nós fizemos as seguintes alterações ao modelo proposto pela pesquisadora:

- a) Junção ou amálgama dos determinantes em primeira posição adjacentes e não adjacentes ao núcleo em um único fator: determinante em primeira posição à esquerda do núcleo.
- b) Junção ou amálgama dos determinantes em segunda posição adjacentes e não adjacentes ao núcleo: determinante em segunda posição à esquerda do núcleo.
- c) Junção ou amálgama dos núcleos de primeira e segunda posição em apenas um fator: núcleo em primeira e segunda posição.

A justificativa para tais amálgamas se deve ao grande número de *knock outs* encontrados por nós ao longo das rodadas nos seguintes fatores: primeira posição não adjacente ao núcleo, núcleo em primeira posição e segunda posição não adjacente ao núcleo. A presença de tais entraves impediam que o GoldVarb 3.0b3 tivesse condições de executar *Binominal Up and Down*.

Vejamos o que os dados revelam:

Tabela 8: Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável *posição do constituinte com referência ao núcleo*

Posição do constituinte com referência ao núcleo	Português culto	P.R.
Determinante em primeira posição à esquerda do núcleo	892/894 99,8%	0.47
Núcleo em primeira e segunda posição	785/969 81%	0.43
Determinante em segunda posição à esquerda do núcleo	77/84 91,7%	0.71
Núcleo em terceira posição em diante	100/116 86,2%	0.81
Modificador à direita do núcleo	128 / 142 90,1%	0.71

Em termos percentuais, nossos resultados não divergem tanto dos encontrados por Guimarães (2014) em sua análise do PPVC. Conforme tabela 8, relativo ao PCVC, obtivemos um índice de marcação de plural de 99,8% e 91,7% para os determinantes em primeira e segunda posição à esquerda do núcleo, respectivamente. No PPVC, Guimarães (2014) obteve, em sua análise, 99,3% para os determinantes em primeira posição imediatamente à esquerda do núcleo e 97,8% para os determinantes em primeira posição não adjacente ao núcleo.

Em relação aos determinantes em segunda posição à esquerda do núcleo, o percentual por ela encontrado foi de 82,3%. Entretanto, quando traduzimos esses resultados em pesos relativos, constatamos logo que nossa hipótese não foi confirmada, completamente, pois esperávamos que os determinantes em primeira posição à esquerda do núcleo recebessem mais marcas de plural. E os dados dos falantes da norma culta não refletiram, probabilisticamente, a nossa expectativa, conforme tabela 8. Enquanto que no PPVC, Guimarães (2014) encontrou para os determinantes em primeira posição imediatamente à esquerda do núcleo peso relativo de 0.84 e 0.88 para os determinantes em primeira posição não adjacente ao núcleo. Nos dados do PCVC encontramos, para esse fator, peso de 0.47.

Nos dados de Lopes (2001) relativo ao português universitário, ao contrário, os elementos não nucleares à esquerda do núcleo, adjacentes e não adjacentes, receberam um índice alto de marcação de plural com pesos de 0.80 e 0.76, respectivamente. Enquanto que nos elementos não nucleares à direita do núcleo o índice de pluralidade foi muito baixo 0.21. Resultado muito distante dos nossos que para esse fator obtivemos 0.71. Quanto aos elementos nucleares, a autora encontrou os seguintes pesos relativos: primeira posição 0.43, segunda 0.28 e terceira 0.14. Em nossos dados, juntamos os núcleos de primeira e segunda posição em um único fator e obtivemos 0.43. Já os núcleos em terceira posição foram os que mais favoreceram a aplicação da regra formal de plural, com peso de 0.81.

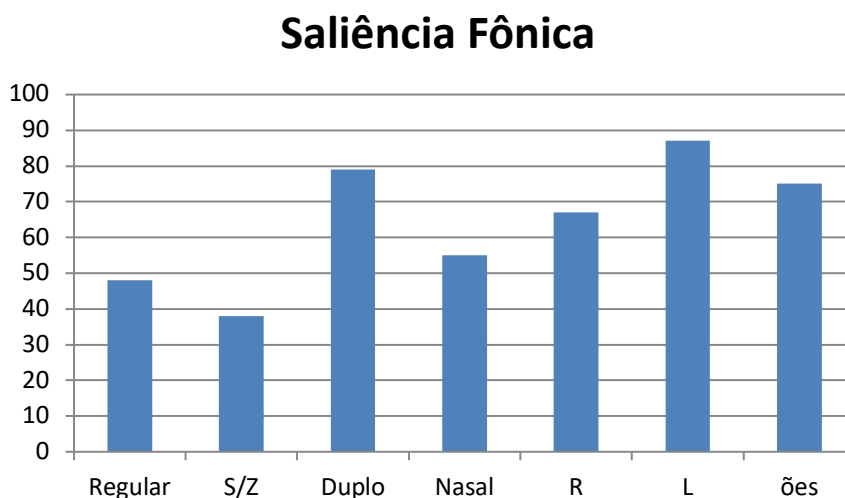
A força condicionadora que tem a posição à esquerda do núcleo na ocorrência de plural nos elementos do SN foi atestada também por Silva (2011). O resultado probabilístico encontrado por ela foi de 0.81 para os constituintes pré-nucleares. Os dados de Martins (2013) atribuíram, do mesmo modo, peso significativo de 0.78 a posição dos elementos não nucleares antepostos ao núcleo. Para Scherre (1998), “a posição à esquerda – a posição de prominência tópica – favorece mais marcas explícitas de plural do que a posição à direita [...], independentemente do nível de escolarização dos falantes” (SCHERRE, 1998, p. 10).

4.1.3 Saliência fônica

Nossa hipótese é a de que os itens lexicais com formação de plural irregular favoreçam mais a indicação de pluralidade por serem mais perceptíveis na relação singular/plural.

Tabela 9: Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável *saliência fônica*

Saliência fônica	Português culto	P.R.
Itens regulares (casa/casas)	1686 / 1870 90,2%	0.48
Itens em –s e –z (ex.: freguês/fregueses; vez/vezes)	44 / 61 72,1%	0.38
Plural duplo (ex.: ovo/ovos; novo/novos)	14 / 15 93,3%	0.79
itens em –m/-em/-ã/-um/ao (ex.: tom/tons)	110 / 119 92,4%	0.55
Itens em –r (ex.: cantor/cantores)	50 / 56 89,3%	0.67
Itens em –l (ex.: animal/animais)	38 / 39 97,4%	0.87
Itens em –ões (ex.: leão/leões)	40 / 45 88,9%	0.75

Gráfico 4: variável saliência fônica no PCVC

Vejamos os dados comparando ao estudo elaborado por Guimarães (2014).

Tabela 10: Frequência e probabilidade da variante “marcas de plural nos constituintes do SN”, segundo a variável *saliência fônica*

Saliência fônica	PPVC	P.R.	Português culto	P.R.
Itens regulares (casa/casas)	1512/2590 58,4%	0.49	1686 / 1870 90,2%	0.48
Itens em –s e –z (ex.: freguês/fregueses; vez/vezes)	70/126 55,6%	0.79	44 / 61 72,1%	0.38
Plural duplo (ex.: ovo/ovos; novo/novos)	4/9 44,4%	0.64	14 / 15 93,3%	0.79
itens em –m/-em/-ã/-um/ao (ex.: tom/tons)	95/166 57,2%	0.29	110 / 119 92,4%	0.55
Itens em –r (ex.: cantor/cantores)	14/46 30,4%	0.53	50 / 56 89,3%	0.67
Itens em –l (ex.: animal/animais)	10/25 40%	0.68	38 / 39 97,4%	0.87
Itens em –ões (ex.: leão/leões)	3/17 17%	0.36	40/45 88,9%	0.75

O princípio da saliência fônica foi outra variável selecionada pelo programa cujo resultado refletiu, no caso do português culto, o que já prevíamos, ou seja, os itens lexicais com formação de plural irregular favorecem mais a indicação de pluralidade por serem mais perceptíveis na relação singular / plural. No entanto, em conformidade com os estudos de Scherre (1988), verificamos, na análise dos dados, certa hierarquia na escala da saliência em que os itens mais marcados foram: itens em –l (0.87), plural duplo (0.79), itens em –ões (0.75), itens em –r (0.67), itens em –m/-em/-ã/-um/-ao (0.55). Os itens regulares (0.48)

superaram apenas os itens em *-s* e *-z* (0.38). Curiosamente, no PPVC, esse último fator foi o que mais favoreceu a marcação de plural, com peso relativo de 0.79. Houve, igualmente, maior favorecimento nos casos de alteração da qualidade vocálica, como em *animal/animais*, mais conhecidos como itens em *-l*, com peso de 0.68, seguido do plural duplo com 0.64. No geral, é possível afirmar que nossa hipótese é confirmada, já que esperávamos que os itens irregulares fossem mais marcados do que os regulares no PCVC.

Comparando nossos dados com os de Lopes (2001), que analisou também o efeito da saliência fônica sobre a concordância nominal no português universitário, notamos um reflexo similar do efeito da saliência fônica em relação a alguns fatores envolvidos nas análises de ambos os estudos. Assim, no PCVC o fator que mais favoreceu a aplicação da regra de plural foram os itens em *-l*, com peso relativo de 0.87. No português universitário de Salvador esse mesmo fator teve marcação categórica de 100%, não sofrendo variação. Comparando ainda os nossos dados com os da autora, em relação a alguns outros fatores da saliência fônica, observamos que no português culto de Vitória da Conquista foram mais marcados também o plural duplo 0.79, os itens em *-ões* 0.75 e em *-r* 0.67. Enquanto que nos dados da capital baiana, entre os universitários, houve uma ligeira diferença na hierarquia desses fatores uma vez que os itens que mais receberam marcas formais foram os itens em *-ões* 0.87, seguidos dos itens em *-r* 0.81, itens em *-s* e *-z* 0.71 e plural duplo 0.65. Em síntese, a pesquisa da autora evidencia, do mesmo modo, que os itens irregulares com maior grau de saliência tendem a ser mais marcados formalmente. Essa foi a mesma conclusão a que chegaram H. Carvalho (1997) e Silva (2011), por exemplo, cujas pesquisas envolveram falantes com diferentes níveis de escolaridade, inclusive o nível superior.

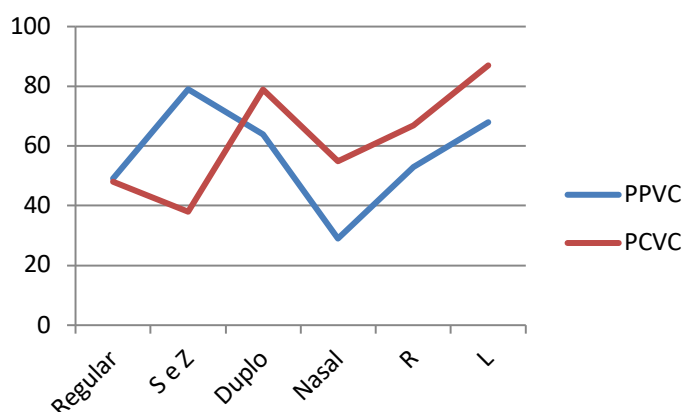
Por sinal, para essa variável, H. Carvalho (1997) chegou aos seguintes resultados probabilísticos: itens irregulares em *-s* e *-z* 0.95, plural duplo e itens em *-l* 0.88, itens em *-r* 0.86 e itens em *-ões* 0.70.

Essa mesma tendência de favorecimento dos itens irregulares na indicação de plural no SN foi encontrada por Silva (2011). Em seu estudo, a autora analisou a fala de informantes de Vitória ES com níveis de escolaridade fundamental, médio e superior. Seus dados demonstraram que os itens lexicais mais salientes tendem a receber mais marcas formais de plural, com peso de 0.61, bem acima do peso de 0.47 para os itens menos salientes.

Campos e Rodrigues (2002), por sua vez, analisaram dados de informantes com formação em nível superior completo ou em andamento, revelando, no estudo de ambas, que a saliência fônica não exerceu nenhuma influência na indicação de pluralidade nos elementos

do SN. As autoras trabalharam com o *corpus* do projeto NURC, o português culto de cinco capitais brasileiras: Porto Alegre, Rio de Janeiro, São Paulo, Salvador e Recife.

Gráfico 5: saliência fônica no PPVC e no PCVC



4.1.4 Classe gramatical do constituinte

No que se refere aos dados do português culto, vale destacar que a variável classe gramatical não foi selecionada pelo programa. No entanto, os resultados do arquivo de células já permite, em parte, uma análise interpretativa dos dados, pois ele informa a quantidade de ocorrências e os percentuais para cada variante em relação a cada fator considerado nos grupos de fatores, como bem observa a professora Josane Oliveira⁴⁴. Lembrando que para o PPVC, essa variável foi selecionada, apresentando pesos relativos que apontam os fatores que mais influenciam a concordância nominal de número. Mas para fins comparativos dos dados, nos reservamos tão somente a apresentar os percentuais de cada uma das normas: popular e culta, conforme tabela 11. Nossa hipótese é a de que os pronomes determinantes e os artigos recebem mais marcas de plural.

Tabela 11: Frequência e percentual da variante “marcas de plural nos constituintes do SN”, segundo a variável *classe gramatical*

Classe gramatical	PPVC	%	Português culto	%
Pronome possessivo	133/145	91,7%	250/254	98,4%
Substantivo	455/1597	28,5%	882/1082	81,5%

⁴⁴ Ministrante do minicurso Introdução ao GoldVarb X: análise sociolinguística, realizado na 25ª Jornada Nacional do GELNE: Grupo de Estudos Linguísticos e Literários do Nordeste, em outubro de 2014.

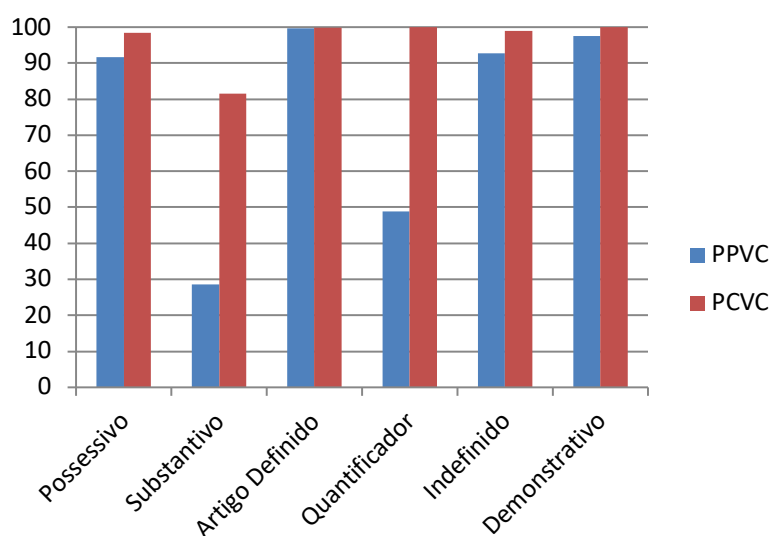
Artigo definido	741 / 743	99,7%	441 / 442	99,8%
Quantificador	23 / 47	48,9%	17 / 17	100%
Pronome indefinido	89 / 96	92,7%	183 / 185	98,9%
Pronome demonstrativo	121 / 124	97,6%	50 / 50	100%
Adjetivos	15 / 92	16,3%	141 / 159	88,7%
Artigo indefinido	129 / 131	98,5%	14 / 14	100%

Alguns resultados observados na tabela 11 evidenciam uma situação contrastiva envolvendo os falantes da norma popular e culta no que tange à marcação de plural em alguns fatores. Foi o que ocorreu com os substantivos, os quantificadores e os adjetivos cujos índices de marcas de plural entre falantes do PCVC foram muito maiores do que os encontrados no PPVC. Todavia, nossa hipótese foi confirmada, em termos percentuais, pois esperávamos que os artigos e os pronomes, que atuam como determinantes, favorecessem a aplicação da regra formal, como de fato ocorreram nas duas normas do português de Vitória da Conquista.

O que pudemos perceber com esses dados é que mesmo em sintagmas nominais com três ou mais constituintes em que os pronomes adjacentes aos núcleos encontram-se pospostos aos artigos, houve também uma tendência de retenção da marca de plural no segundo elemento. Exemplo: *com os meus irmão*. O que pode ser explicado pelo processamento com paralelismo proposto por Scherre (1988), cujo desdobramento na concordância nominal de número se dá com base no princípio de que a ocorrência de marcas de plural em um item anterior leva à marcas no item seguinte. Vale lembrar que o alto índice da variante explícita de plural nos artigos definidos e indefinidos, no PPVC e no PCVC, está relacionado, certamente, a sua posição no sintagma, que é normalmente a primeira.

Essa é a mesma constatação a que chegou Lopes (2001). Em sua análise estatística, a pesquisadora demonstrou que os adjetivos e os artigos foram os fatores com maior peso relativo (0.80 e 0.73, respectivamente). Os pronomes possessivos com 0.67 e os indefinidos com 0.58 também foram bem marcados formalmente. Quanto aos pronomes demonstrativos, esse fator não exibiu variação, assim como aconteceu com nossos dados, atingindo um percentual de 100%. Os menores pesos relativos encontrados pela autora foram os dos substantivos (0.35) e quantificadores (0.24).

Gráfico 6: Marcas de plural nos constituintes do SN, segundo a variável classe gramatical.



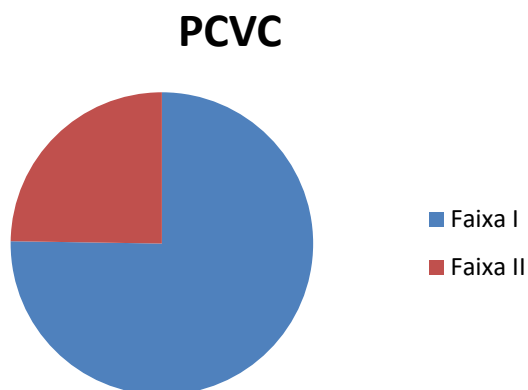
4.1.5 Variável social: *faixa etária*

Em relação à idade, nossa análise recai sobre dados de falantes de duas faixas etárias: jovens de 20 a 35 anos (faixa I) e adultos de 36 a 50 anos (faixa II). Para essa variável, partimos da hipótese de que os adultos fazem mais uso da variante explícita de plural nos elementos do SN.

Tabela 12 – A aplicação da regra de concordância nominal de número na variável faixa etária

	PCVC	P.R.
Faixa I	1058 / 1086 97,4%	0.76
Faixa II	924 / 1119 82,6%	0.25

Conforme se pode observar na tabela 12, nossa hipótese não foi atestada uma vez que os jovens tendem a apresentar mais marcas de concordância nominal de número no português culto de Vitória da Conquista.

Gráfico 7: Variável faixa etária

Vamos comparar os nossos dados aos de Guimarães (2014), conforme tabela 13.

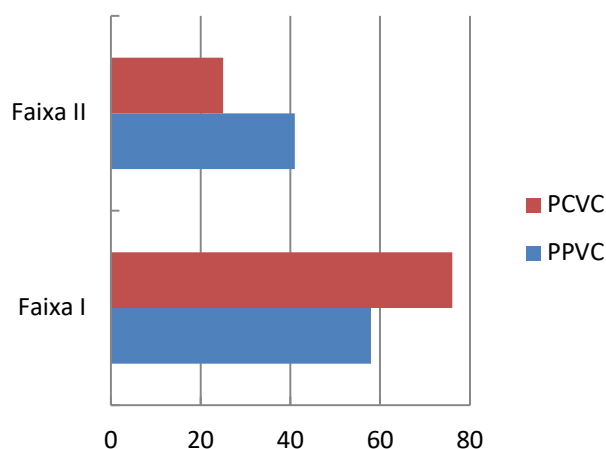
Tabela 13 – A aplicação da regra de concordância nominal de número na variável faixa etária

	PPVC	P.R.	PCVC	P.R.
Faixa I	499/837 59,6%	0.58	1058 / 1086 97,4%	0.76
Faixa II	683 / 1213 56,3%	0.41	924 / 1119 82,6%	0.25

Assim como ocorreu no português culto, no português popular os jovens lideram o uso da variante padrão. O peso relativo de 0.58 no PPVC sugere-nos uma provável mudança em curso em direção a um sistema com concordância. Segundo Guimarães (2014),

[...] o processo de urbanização e o crescimento econômico estariam configurando alterações no PPVC, impulsionado pelos mais jovens, por estarem mais sujeitos a valores de outros grupos, bem como por se colocarem próximo a um ideal de urbanidade (GUIMARÃES, 2014).

Já no português culto, o peso relativo de 0.76 parece evidenciar uma situação de variação estável, sobretudo pelo baixo índice de variabilidade nos constituintes sintagmáticos. Além disso, a grande maioria dos jovens falantes do PCVC tem nível superior completo ou em andamento, ao contrário da maioria dos adultos cuja formação é em nível técnico profissional.

Gráfico 8: Variável faixa etária no português popular e culto

De modo geral, os nossos resultados estão em consonância com aqueles encontrados por H. Carvalho (1997), Andrade (2003), Silva (2011), entre outros.

Assim, em sua pesquisa, H. Carvalho (1997) encontrou peso relativo de 0.61 para a primeira faixa, revelando-nos que os mais jovens são mais favorecedores do uso da concordância nominal.

Os resultados de Andrade (2003) seguem essa mesma linha de tendência de maior índice de plural no SN entre os mais jovens. O peso de 0.66 para a faixa I, bem acima de 0.46 para a faixa II, leva a autora a reconhecer que “o dialeto de Helvécia está passando por um processo de mudança em curso” (ANDRADE, 2003, P. 125).

Os dados de Silva (2011) também revelaram maior favorecimento da concordância nominal entre os mais jovens, com peso relativo de 0.61.

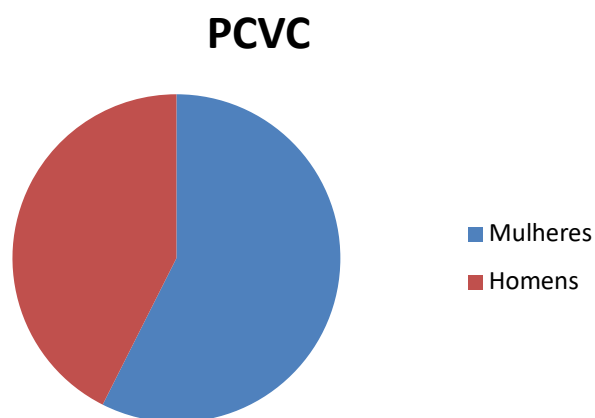
4.1.6 Variável social: *sexo*

Com base nos resultados encontrados por Scherre (1988), Santos (2010) e Martins (2013), segundo os quais as mulheres fazem mais concordância nominal de número no SN, elaboramos nossa hipótese para a variável sexo, esperando, igualmente, que na comunidade de fala de Vitória da Conquista as mulheres façam mais o uso de marcas formais de plural do que os homens.

Tabela 14 – A aplicação da regra de concordância nominal de número na variável sexo

	Português culto	P.R.
Mulheres	948 / 1006 94,2%	0.58
Homens	1034 / 1199 86,2%	0.43
Total	1982 / 2.205	

No que tange à variável social sexo, selecionada pelo programa, verificamos que no PCVC há uma maior tendência de aplicação da regra formal entre as mulheres, com peso probabilístico de 0.58. Esse resultado comprova nossa hipótese.

Gráfico 9: A aplicação da regra de concordância nominal de número na variável sexo

Para fins comparativos, apresentamos na tabela 15 os dados da norma popular e culta relativos à variável sexo.

Tabela 15 – A aplicação da regra de concordância nominal de número na variável sexo

	PPVC	P.R.	Português culto	P.R.
Mulheres	940/1539 61,1%	0.57	948 / 1006 94,2%	0.58
Homens	768/1440 53,3%	0.41	1034 / 1199 86,2%	0.43
Total	1708 / 2979		1982 / 2.205	

Os resultados da tabela 15 mostram-nos que no português popular, assim como aconteceu no português culto, são as mulheres que fazem mais uso do plural. O que tanto pode sugerir que elas sejam muito mais atentas ao uso das formas mais prestigiadas linguisticamente, como também que elas, apesar das grandes mudanças comportamentais das últimas décadas, sejam mais cerceadas e mais cobradas para ter um comportamento mais em

conformidade com as convenções sociais do que os homens. Para Guimarães (2014), em relação aos dados do PPVC, “os resultados demonstram que a “emancipação feminina” com a ampliação de seu espaço de atuação favorece à aquisição de padrões linguísticos cada vez mais próximos da norma de prestígio” (GUIMARÃES, 2014). Esse argumento ganha ainda mais força quando o direcionamos às informantes do PCVC que além de terem seu espaço de atuação social mais amplo, estão inseridas no contexto da cultura letrada, tendo um contato muito mais constante e ativo com as formas linguísticas mais prestigiadas socialmente.

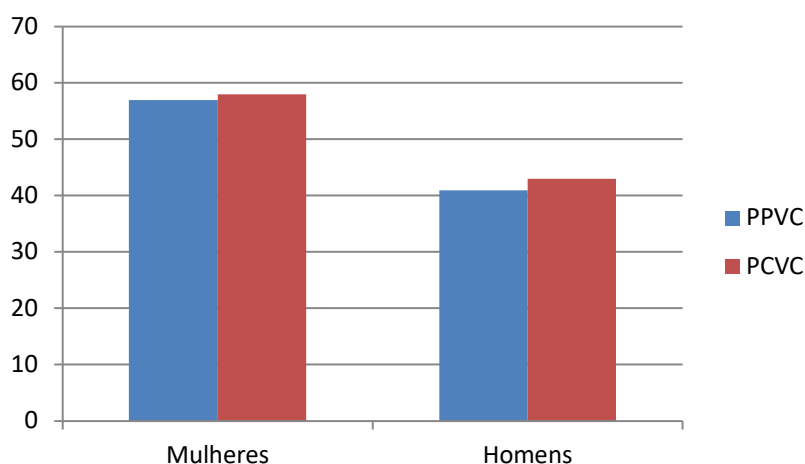
Em relação aos homens, uma explicação para a menor incidência de marcas de plural nos constituintes do SN pode ser atribuída ao fato de eles, por uma questão cultural, serem menos reprimidos no seu comportamento e, conseqüentemente, se permitirem a adoção de uma fala muito mais relaxada e espontânea.

Com isso, eles acabam liderando as mudanças para variantes não padrão enquanto que as mulheres se mostram mais sensíveis à variante padrão. Ressaltamos que nos dados do PCVC essa variável foi selecionada em quinto lugar.

Outros estudos, como o de Scherre (1988), mostraram, também, que as mulheres usam mais o plural nos constituintes flexionáveis do SN, com peso relativo de 0.58 contra 0.42 dos homens.

A mesma tendência de uso da forma padrão entre o público feminino foi demonstrada por Martins (2013) que encontrou, probabilisticamente, peso relativo de 0.52 para as mulheres e 0.47 para os homens.

Gráfico 10: A aplicação da regra de concordância nominal de número na variável sexo no PPVC e PCVC



Vale lembrar que algumas pesquisas como as apresentadas por Rodrigues⁴⁵, bem como os estudos realizados por Bortoni-Ricardo (2005, p. 91-102), no início da década de 1980, em Brasilândia, cidade periférica de Brasília e os estudos feitos na zona rural, no interior do estado da Bahia, sob liderança de Lucchesi, têm revelado que são os homens que usam mais a variante padrão de prestígio e não as mulheres. Ora, esses resultados encontrados em tais comunidades nos levam a refletir acerca da necessidade de evitarmos uma mera interpretação mecânica do fenômeno da variação envolvendo as variáveis estruturais e sociais a fim de evitarmos, igualmente, algumas generalizações um tanto precipitadas, como, por exemplo, a de que as mulheres fazem mais concordância do que os homens.

É preciso considerar, antes de tudo, o contexto social em que se encontra a população pobre de um país do terceiro mundo, ainda em desenvolvimento, não plenamente industrializado, como é o caso do Brasil. Por conta dessa realidade, é possível concluir que nessa população das periferias das grandes cidades e da zona rural são os homens que saem mais da comunidade, estão mais inseridos no mercado de trabalho, enquanto que as mulheres ficam mais em casa. E como a mudança no português brasileiro popular vem de cima para baixo, de fora para dentro da comunidade, quem tem mais contato externo é que vai liderar essa mudança de assimilação da forma de prestígio. E nesse caso são os homens.

4.1.7 Variável social: *escolaridade*

A hipótese por nós levantada é a de que a aplicação da regra de concordância nominal é maior entre os falantes com grau de escolaridade em nível superior uma vez que eles estão muito mais expostos a atividades culturais diversificadas nas quais predominam o uso das formas linguísticas mais prestigiadas.

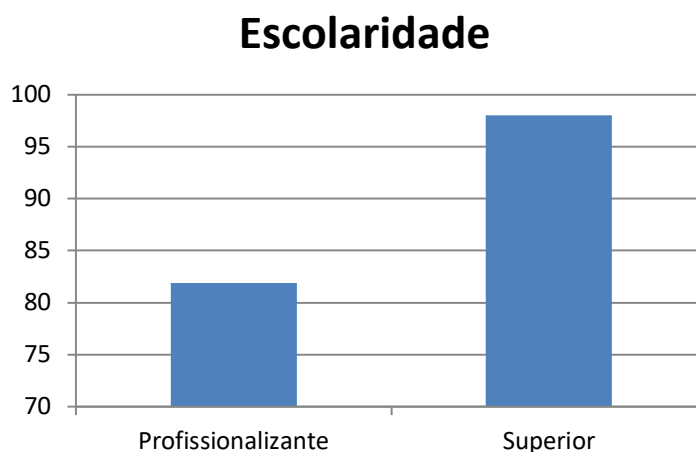
A fim de controlarmos essa variável dividimos os informantes do PCVC em dois grupos: os de nível superior e os de formação em cursos profissionalizantes. A lógica para essa divisão se pautou no fato dos cursos de nível superior terem a chance de incluir em seus currículos matérias/disciplinas de cunho humanístico o que acarreta a leitura de textos mais densos e a necessidade de produção acadêmica, além de exposições em seminários, congressos, simpósios, ou seja, atividades que requerem o uso de norma culta e o cuidado ao falar em público. Os resultados obtidos foram os seguintes:

⁴⁵ Disponível em: <https://periodicos.ufsc.br/index.php/forum/article/>.

Tabela 16: Variável escolaridade

Escolaridade	Português culto	P.R.
+ de 11 anos: curso profissionalizante	911 / 1112 81,9%	0.26
+ de 11 anos: nível superior	1071 / 1093 98%	0.74

Observando a tabela 16, percebemos logo que nossa hipótese se confirmou, já que, probabilisticamente, os resultados mostram que a parcela de falantes do português culto, com formação em nível superior, tende a fazer mais concordância nominal de número, com peso relativo de 0.74. O que pode ser explicado, conforme hipótese por nós levantada, pela exposição mais constante dos estudantes a atividades culturais diversificadas nas quais há a preponderância do uso das normas culta e padrão da língua portuguesa.

Gráfico 11: variável escolaridade no PCVC

Vale lembrar que, no português culto, essa variável foi selecionada em segundo lugar dentre as variáveis controladas em nossa pesquisa.

Na tabela 17, apresentamos para essa variável os resultados do português popular analisado por Guimarães (2014).

Tabela 17 - A aplicação da regra de concordância nominal de número na variável escolaridade

Escolaridade	PPVC	P.R.
1 a 2 anos	896 / 1589 56%	0.44
3 a 4 anos	255 / 458 55,7%	0.49
5 anos	557 / 932 59,8%	0.59

Observando os resultados do português popular e do português culto, percebemos logo o que já é quase um consenso e um truísmo entre os pesquisadores, a saber, quanto maior o nível de escolaridade do falante, maior o índice de marcas de plural.

É importante destacar também que enquanto na norma culta o falante com nível superior completo faz muito mais concordância nominal, no português popular a frequência de aplicação da regra de concordância é bem menor. Essa diferença entre o comportamento objetivo dos falantes pode refletir, muitas vezes, nas diferenças dos juízos de avaliação que tendem a separar, socialmente, os falantes da norma culta dos falantes da norma popular, constituindo, nesse caso, a base do preconceito linguístico que ainda impera na sociedade brasileira como um todo.

CONCLUSÃO

Investigamos, nessa pesquisa, amostras de fala de informantes da zona urbana de Vitória da Conquista, classificados, conforme os anos de formação escolar, em falantes do português popular (até cinco anos de escolarização) e falantes do português culto (com mais de onze anos de escolarização), considerando, nesse último caso, a educação profissional técnica e o nível superior.

Os dados do português popular foram analisados por Guimarães (2014) em sua pesquisa, cujos resultados utilizamos em nosso estudo para comparar com os resultados da análise que fizemos do português culto, no que se refere à variação na concordância nominal de número no SN.

Em nossa análise com os dados do português culto, partimos de uma perspectiva mórfica ou atomística na qual cada constituinte flexionável do sintagma nominal se constituiu em unidade de análise. As variáveis selecionadas pelo programa Goldvarb foram, em primeiro lugar, posição linear do constituinte; em segundo lugar, escolaridade; em terceiro lugar, faixa etária; em quarto lugar, posição do constituinte com referência ao núcleo; em quinto lugar, sexo e, em sexto lugar, saliência fônica. A variável classe gramatical do constituinte não foi selecionada em nossa amostra.

A hipótese para a variável posição linear do constituinte foi atestada, pois os dados confirmaram o que já prevíamos, ou seja, a primeira posição foi a mais favorável a marcas de plural, assim como ocorreu no português popular.

Em relação à variável social escolaridade, nossa hipótese se confirmou, tendo em vista que os resultados dos pesos relativos mostraram-nos que a parcela de falantes do português culto, com formação em nível superior, tende a fazer muito mais concordância nominal de número do que os falantes com formação em curso profissionalizante.

No que se refere à faixa etária, nossa hipótese não foi atestada uma vez que foram os jovens, e não os adultos, que apresentaram mais marcas de concordância nominal de número, inclusive, nas duas vertentes do português de Vitória da Conquista.

Em nossa hipótese para a variável posição do constituinte com referência ao núcleo, esperávamos que os determinantes em primeira posição à esquerda do núcleo recebessem mais marcas de plural no português culto, assim como ocorreu no português popular. O que não se confirmou, se considerarmos tão somente o peso relativo. Porém, o percentual de quase 100%, mais precisamente de 99,8%, revela-nos a força que tem esse fator na retenção da marca de plural, sugerindo-nos, com isso, a confirmação de nossa hipótese.

A variável sexo, por sua vez, atestou nossa hipótese segundo a qual as mulheres fariam mais uso da variante explícita de plural nos constituintes sintagmáticos do que os homens. O que de fato se confirmou em nosso estudo.

Foi confirmada, do mesmo modo, a nossa hipótese para a variável saliência fônica, pois os itens lexicais com formação de plural irregular favoreceram, na sua grande maioria, mais a indicação de pluralidade por serem mais perceptíveis na relação singular/plural.

De modo geral, nosso estudo evidenciou um altíssimo índice percentual da variante explícita de plural entre os informantes do português culto. Com exceção de um único fator da variável estrutural saliência fônica - os itens em *-s* e *-z* que favoreceram a marca de plural em 72,1% - todos os demais fatores de todas as variáveis independentes atingiram percentuais acima de 80%. O que comprova que, em termos percentuais, os falantes da norma culta do português de Vitória da Conquista fazem muito mais concordância nominal de número do que os falantes do português popular. Essa constatação só faz reforçar o que muitas outras pesquisas já comprovaram, a saber: quanto maior o tempo de exposição à escola e o nível de escolarização, maior o uso de formas linguísticas mais prestigiadas, como a indicação de plural nos constituintes flexionáveis do SN.

Por outro lado, os pesos relativos observados nos dados do PPVC e do PCVC revelaram uma certa oscilação em relação a alguns fatores analisados em ambos os domínios. Enquanto que para a variável posição do constituinte com referência ao núcleo, o determinante em primeira posição, adjacente e não adjacente ao núcleo, favoreceu a marcação formal de plural no português popular; no português culto, essa tendência não foi demonstrada, contrariando o alto percentual encontrado para tal fator, 99,8%.

No mais, os resultados não deixam dúvida de que o fenômeno da variabilidade na concordância nominal de número no português de Vitória da Conquista se apresenta, nitidamente, como um caso de variação estável em que as duas variantes: presença e ausência de marcas de plural no SN coexistem na fala dos informantes. No entanto, o índice de variação foi bem maior entre os falantes do português popular, havendo, inclusive, um processo de mudança em curso em direção à aquisição de marcas de plural entre os mais jovens.

Com esse estudo, esperamos contribuir significativamente para o conhecimento da realidade sociolinguística de Vitória da Conquista.

REFERÊNCIAS

- ANDRADE, Patrícia Ribeiro de. **Um fragmento da constituição sócio-histórica do português do Brasil**: variação na concordância nominal de número em um dialeto afro-brasileiro. Dissertação de mestrado, UFBA, Instituto de Letras, 2003.
- ANTONINO, Vivian. **Português popular de Salvador**: uma análise da concordância nominal em predicativos do sujeito e em estruturas passivas. Tese de doutorado, UFBA, Instituto de Letras, 2012.
- BAGNO, Marcos. **Nada na língua é por acaso**: por uma pedagogia da variação linguística. São Paulo: Parábola Editorial, 2007.
- BENVENISTE, Émile. **Problemas de linguística geral I**. 5ª edição, Campinas, SP: Pontes Editores, 2005.
- BORTONI-RICARDO, Stella Maris. **Nós chegemos na escola, e agora?**: Sociolinguística e educação. São Paulo: Parábola, 2005.
- BRANDÃO, Sílvia Figueiredo. **Concordância nominal em duas variedades do português**: convergências e divergências. Revista Veredas. Atemática, 2011.
- BUENO, Elza S. da Silva. **Influências de variáveis relacionadas na concordância de número em português**. In: 50º Seminário do Grupo de Estudos Linguísticos do Estado de São Paulo, 2002.
- CALVET, Louis-Jean. **Sociolinguística**: uma introdução crítica. São Paulo: Parábola Editorial, 2002.
- CÂMARA JR, Joaquim Mattoso. **História e estrutura da língua portuguesa**. Rio de Janeiro: Padrão, 1975.
- CAMPOS, Odette G. L. de Souza; RODRIGUES, Ângela C. S. **Flexão Nominal**: indicação de pluralidade no sintagma nominal. In: ILARI, Rodolfo (org.). Gramática do português falado. V. II: níveis de análise linguística. 4ª ed. Campinas: editora da UNICAMP, 2002.
- CARVALHO, Hebe. **Concordância nominal**: uma análise variacionista. João Pessoa: UFPB, 1997.
- CARVALHO, Raimunda Coelho de. **A concordância de número no sintagma nominal na fala urbana de Rio Branco**. Mestrado em Linguística. Campinas: UNICAMP, 1997.
- CASTILHO, Ataliba T. de. **Fundamentos teóricos da gramática do português culto falado no Brasil**: sobre o segundo volume, classes de palavras e as construções gramaticais. Alfa, São Paulo, v. 51, n.1, p. 99-135, 2007.
- CASTRO, Flávio Marcelo Bueno de; PEREIRA, Vinícius Carvalho. **A concordância nominal na norma culta em Cuiabá**. Revistas LetrasMil. V. 1, n. 3, 2012.

CUNHA, Celso. **A questão da norma culta brasileira**. Rio de Janeiro: Tempo Brasileiro, 1985.

FARIAS, Roberto Carlos. **A atual situação diglósica do Brasil**: suas origens no bilinguismo dos primeiros séculos de colonização. In: Actas do Primeiro Simpósio Internacional sobre Bilinguismo. Vigo, Galiza, 1997. Disponível em: <http://webs.uvigo.es/ssl/actas1997/06/Farias.pdf>. Acesso em 10 de outubro de 2014.

FERNANDES, Marisa. **Concordância nominal na região sul**. Dissertação de mestrado. Florianópolis: UFSC, 1996.

FONTES, José Raimundo. **O Novo Ciclo de Desenvolvimento de Vitória da Conquista-BA**, 2009. Disponível em: <<http://www.conquistadetodos.com.br/artigo>> Acesso em 09 de abril de 2014.

GUIMARÃES, Maria Aparecida Souza. **Concordância nominal de número no português popular do Brasil**: estudo de variação e mudança no vernáculo conquistense. Dissertação de mestrado. Programa de Pós-Graduação em Linguística, UESB, 2014.

GUY, Gregory. **As comunidades de fala**: fronteiras internas e externas. Abralin, 2001. Disponível em <http://sw.npd.ufc.br/abralin/anais_con2int_conf02.pdf>

ILARI, Rodolfo; BASSO, Renato. **O português da gente**: a língua que estudamos/a língua que falamos. São Paulo: Contexto, 2009.

JAKOBSON, Roman. **Linguística e comunicação**. São Paulo, Cultrix, 1969.

LABOV, William. **Padrões sociolinguísticos**. Tradução Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso. São Paulo: Parábola Editorial, 2008.

LOBO, Tânia. **A formação histórica do português brasileiro**: o estado da questão. Comunicação ao XI congresso da ALFAL. Gran Canária, 1996.

LOPES, Norma. **A concordância nominal, contexto linguístico e sociedade**. Tese de doutorado, UFBA, Faculdade de Letras, 2001.

LUCCHESI, Dante. **Variação e norma**: elementos para uma caracterização sociolinguística do português do Brasil. In: *Revista Internacional de Língua Portuguesa*. n. 12, 1994.

LUCCHESI, Dante. **A variação na concordância de gênero em uma comunidade de fala afro-brasileira**: novos elementos sobre a formação do português popular do Brasil. Rio de Janeiro: UFRJ, Faculdade de Letras. Tese de Doutorado em Linguística: 2000.

LUCCHESI, Dante. **As duas grandes vertentes da história sociolinguística do Brasil (1500-2000)**. D.E.L.T.A., 17:1, 2001.

LUCCHESI, Dante. O conceito de transmissão linguística irregular e o processo de formação do português no Brasil. In: RONCARATI, Claudia; ABRAÇADO, Jussara (orgs.). **Português brasileiro**: contato linguístico, heterogeneidade e história. Rio de Janeiro: 7 letras, 2003.

LUCCHESI, Dante. **Sistema, Mudança e Linguagem**. São Paulo: Parábola Editorial, 2004.

LUCCHESI, Dante. **Parâmetros sociolinguísticos do português brasileiro**. Revista da ABRALIN, v. 5, n. 1 e 2, p. 83-112, dez. 2006.

LUCCHESI, Dante; BAXTER, Alan e RIBEIRO, Ilza. (Orgs.) **O português afro-brasileiro**. Salvador: EDUFBA, 2009.

LUCCHESI, Dante; BAXTER, Alan. A transmissão linguística irregular. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (Orgs.). **O português Afro-Brasileiro**. Salvador: Edufba, 2009.

LUCCHESI, Dante. **Os limites da variação e da invariância na estrutura da gramática**. Revista da ABRALIN, v. Eletrônico, n. Especial, p. 227-259. 2ª parte 2011. Disponível em: www.abralin.org/revista/RVE2/6v. Acesso em 20 de junho de 2014.

LUCCHESI, Dante. **A teoria da variação linguística: um balanço crítico**. Estudos Linguísticos, São Paulo, 2012.

MAGALHÃES, Livia Diana Rocha. **Políticas educacionais e trajetórias geracionais: primeiros relatos da pesquisa em Vitória da Conquista-BA**. In: Anais do VI colóquio do Museu Pedagógico. Vitória da Conquista, 2006.

MARCONI, M.A.; LAKATOS, E.M. **Técnicas de Pesquisa**. 2. ed. São Paulo, Atlas, 1990.

MARTINS, Flávia Santos. **Variação na concordância nominal de número na fala dos habitantes do alto Solimões (Amazonas)**. Tese de doutorado. Florianópolis: UFSC, 2013.

MATTOS E SILVA, Rosa Virgínia. **Uma interpretação para a generalizada difusão da língua portuguesa no território brasileiro**. *Gragoatá*. Niterói, n. 9, p. 11-27, 2 sem. 2000.

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

MATTOS E SILVA, Rosa Virgínia. **Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa**. Cadernos de Letras da UFF – Dossiê: Literatura, língua e identidade, no 34, p. 31-41, 2008.

MATTOSO, Kátia de Queirós. **Ser escravo no Brasil**. 3ª ed. São Paulo: Brasiliense, 1990.

MEC. **Parâmetros Curriculares Nacionais: primeiro e segundo ciclos do ensino fundamental: língua portuguesa**. 3ª ed. Brasília: MEC/SEF, 2001.

MOLLICA, Maria Cecília; BRAGA, Maria Luiza (orgs.). **Introdução à Sociolinguística: o tratamento da variação**. 4ª. ed. São Paulo: Contexto, 2010.

NARO, Anthony Julius; SCHERRE, Maria Marta Pereira. **Origens do Português Brasileiro**. São Paulo: Parábola Editorial, 2007.

NOLL, Volker. **O português brasileiro: formação e contrastes**. São Paulo: Globo, 2008.

NOVAIS, Idelma aparecida Ferreira. **Produção e Comércio na Imperial Vila da Conquista (Bahia, 1840-1888)**. Faculdade de Filosofia e Ciências Humanas, Universidade Federal da Bahia, 2008.

RIBEIRO, Ilza. **A origem do português culto. A escolarização**. Comunicação em encontro da UNIFACS. Salvador, 1999.

RODRIGUES, José Honório. **A vitória da língua portuguesa**. Humanidades. Volume 1, número 4. Universidade de Brasília, 1983.

SANTOS, Lília Soares Miranda. **Sobre a ausência de concordância nominal no português falado em Pedro Leopoldo-MG: uma abordagem variacionista**. Dissertação de mestrado. Belo Horizonte: UFMG, 2010.

SANTOS, Ocerlan Ferreira; NASCIMENTO, Washington Santos. **Dimensões da Vida Escrava na Imperial Vila da Vitória nos últimos anos da escravidão (1870-1888)** (2010). Disponível em: <http://periodicos.uesb.br/index.php/politeia/article>. Acesso em: 21 abr. 2014.

SAPIR, Edward. **A linguagem**: introdução ao estudo da fala. Rio de Janeiro: Acadêmica, 1971.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. 3ª ed. São Paulo: Cultrix, 2012.

SCHERRE, Maria Marta Pereira. **Reanálise da Concordância Nominal em Português**. Tese de Doutorado. Rio de Janeiro: UFRJ, 1988.

SCHERRE, Marta. **Aspectos da concordância de número no português do Brasil**. Revista Internacional de Língua Portuguesa (RILP) – Norma e Variação do Português. Associação das Universidades de Língua Portuguesa, 1994.

SCHERRE, Maria Marta Pereira. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: OLIVEIRA E SILVA, G. M. de; SCHERRE, M. M. P. (Org.). **Padrões sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro**. Rio de Janeiro: Tempo Brasileiro; Universidade Federal do Rio Janeiro, Departamento de Linguística e Filologia, 1996.

SCHERRE, M. M. P. & NARO, A. J. **Sobre a concordância de número no português falado do Brasil**. In Ruffino, Giovanni (org.) *Dialettologia, geolinguística, sociolinguística*. (Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza) Centro di Studi Filologici e Linguistici Siciliani, Università di Palermo. Tübingen: Max Niemeyer Verlag, 5:509-523, 1998.

SCHNEIDER, Simone Daise. **Concordância nominal na fala de crianças de 3 a 6 anos de idade do município de Novo Hamburgo: variação linguística na infância**. Tese de Doutorado. Porto Alegre: UFRGS, 2012.

SILVA, Janaína Biancardi de. **A concordância nominal na fala capixaba**. I congresso nacional de estudos linguísticos, Vitória-ES, 2011.

SILVA, Jorge Augusto Alves da. **A concordância verbal de terceira pessoa do plural no português popular do Brasil: um panorama sociolinguístico de três comunidades do interior do estado da Bahia.** Tese de doutorado, UFBA, Instituto de Letras, 2005.

SILVA, Jorge Augusto Alves; SOUSA, Valéria Viana. **Pelo “Sertão da Ressaca”: contribuições para a compreensão da sócio-história do Português Popular do Brasil.** Disponível em: <http://www.revistas.uneb.br/index.php/tabuleirodeletras/article/view/361>. Acesso em: 15 abr. 2014.

SILVA-DE-SOUSA, Maria Aparecida. **A Conquista do Sertão da Ressaca: povoamento e posse da terra no interior da Bahia.** 1998. 172p. Dissertação (Mestrado em História) - Faculdade de Filosofia e Ciências Humanas, Universidade Federal de Minas Gerais, Belo Horizonte, 1998.

SILVA NETO, Serafim da. **Introdução ao estudo da língua portuguesa do Brasil.** 2. ed. Rio de Janeiro: INL, 1963.

SOUSA, Angelita Cunha da Silva. **A Rua do Maga-Sapo: cotidiano e representações da prostituição em Vitória da Conquista-BA (1950-1971).** Programa de Pós-Graduação em Letras: Cultura, Educação e Linguagens, Universidade Estadual do sudoeste da Bahia, 2013.

SOUZA, Daniela Moura Rocha de. **Memória de professores intelectuais como interlocutores do republicanismo em Vitória da Conquista entre os anos de 1910 até 1945.** Dissertação de mestrado. Vitória da Conquista: UESB, 2009.

TANAJURA, Mozart. **História de Conquista: Crônica de uma cidade.** Vitória da Conquista: Brasil-Artes Gráficas, v. 01, 1992.

TAVARES, Luís Henrique Dias. **História da Bahia.** 10. ed. Revista e ampliada. Salvador: EDUFBA, 2001.

VOTRE, Sebastião; RONCARATI, Cláudia (organizadores). **Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica.** Rio de Janeiro: 7Letras, 2008.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. **Fundamentos empíricos para uma teoria da mudança linguística.** São Paulo: Parábola Editorial, 2006.

APÊNDICE

Algumas ocorrências de fala de informantes do português culto em relação à posição linear do constituinte:

(-1PiRAFS	<i>costurá <u>muito</u> bonés</i>
(-1PaRAMP	<i>ouvia a boas histórias</i>
(-2SiRAMS	<i>algumas <u>outra</u> coisa não</i>
(-2SiRAMP	<i>com os outro irmão</i>

ANEXO

Chave de codificação – Análise Mórfica**1. VARIÁVEL DEPENDENTE****Marcação do plural em cada constituinte do SN**

- (+) plural marcado
- (-) plural não marcado

VARIÁVEIS ESTRUTURAIS**2. Posição linear do constituinte**

- (1) primeira posição
- (2) segunda posição
- (3) terceira posição
- (4) quarta posição em diante

3. Posição do constituinte com referência ao núcleo do sn‣ **DETERMINANTE**

(P) primeira posição imediatamente à esquerda do núcleo

Ex.: **Meus** filho

(p) primeira posição não adjacente ao núcleo

Ex.: **Meus** três filho

(S) segunda posição imediatamente à esquerda do núcleo

Ex.: Todos **meus** filho

(s) segunda posição não adjacente ao núcleo

Ex.: Todos **meus** melhor amigo

(T) terceira posição em diante à esquerda do núcleo

Ex.: Todosmeus **melhor** amigo

‣ **NÚCLEO**

(N) núcleo em primeira posição

Ex: Comprei **bananas** madura

(n) núcleo em segunda posição

Ex: Meus **filho**

(t) núcleo em terceira posição em diante

Ex: Os oito **irmão**

‣ **MODIFICADORES**

(A) constituinte imediatamente à direita do núcleo

Ex: Meus irmão **maió**

(U) constituinte à direita do núcleo não adjacente a ele

Ex: Meus filho mais **velho**

4. Classe gramatical do constituinte

- (a) artigo definido
- (u) artigo indefinido

- (d) pronome demonstrativo
- (p) pronome possessivo
- (i) pronome indefinido [muito(s), tanto(s), outro(s), algum(ns)]
- (q) quantificador *todo(s)/tudo*
- (s) substantivo
- (j) adjetivo
- (?) substantivo coletivo

5. Saliência fônica

- (D) plural duplo (ex.: ovo/ovos, novo/novos)
- (l) itens em -l (ex.: animal/animais)
- (~) itens em -ões (ex.: leão/leões)
- (*) itens em -ães (ex.: pão/pães)
- (!) itens em -m/-em/-ã/-um/ão (ex.: tom/tons; armazém/armazens; irmã/irmãs; algum/alguns; irmão/ irmãos)
- (r) itens em -r (ex.: cantor/cantores)
- (\$) itens em -s e -z (ex.: freguês/fregueses; vez/vezes)
- (R) itens regulares (ex.: casa/casas)

VARIÁVEIS SOCIAIS

6. Idade

- (J) 20 a 35 anos
- (A) 36 a 50 anos

7. Sexo

- (F) feminino
- (M) masculino

8. Escolaridade

- (P) + de 11 anos de escolarização: curso profissionalizante
- (S) + de 11 anos de escolarização: nível superior