UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS GRADUAÇÃO EM SEMIÓTICA E LINGUÍSTICA GERAL

FERNANDO GOMES DA SILVA

Alagoanos em São Paulo e a Concordância Nominal de Número

São Paulo
2014

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS GRADUAÇÃO EM SEMIÓTICA E LINGUÍSTICA GERAL

FERNANDO GOMES DA SILVA

Alagoanos em São Paulo e a Concordância Nominal de Número

Dissertação apresentada ao Programa de Pós-Graduação em Semiótica e linguística Geral do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo para a obtenção do título de mestre em Linguística.

Orientador: Prof. Dr. Ronald Beline Mendes

São Paulo
2014

AGRADECIMENTOS

Agradeço à CAPES pelo apoio à pesquisa.

Agradeço ao meu orientador, Prof. Dr. Ronald Beline Mendes, pela oportunidade de desenvolver esta pesquisa que é tão importante para mim, pela amizade, paciência, pelas discussões e inúmeras horas dedicadas à leitura dos meus textos. Agradeço também aos funcionários do Departamento de Linguística pela atenção e disponibilidade em todas as vezes em que precisei de ajuda.

Agradeço à Profa. Dra. Olga Coelho Sanssone por ter participado de todos os momentos importantes na construção dessa pesquisa. À Prof. Dra. Margarida Petter por todas as orientações e conversas sobre o meu trabalho. Ao Prof. Sebastião Gonçalves (UNESP – São José do Rio Preto), por todos os comentários e sugestões para o aperfeiçoamento dessa pesquisa.

Agradeço aos amigos do GESOL-USP, pela amizade, pelas risadas, pelos desabafos, pelas sugestões construtivas e por fazerem parte da minha vida: Marília, Mariane, Wendel, Livia, Dayane, Rafael Ciano, Rafael Rocha e Larissa.

Agradeço imensamente aos meus pais, Adão e Josefa, pelo amor incondicional, pela força, pelo apoio emocional e por sempre acreditarem e apostarem todas as suas fichas em mim. Agradeço também a todos os meus familiares (primos (as) e tios (as)) que contribuíram para a minha formação como ser humano.

Agradeço ao meu amigo e companheiro (de todas as horas), Júnior, pela ajuda emocional e espiritual, pelas longas e confortáveis conversas, por me apoiar em todas as minhas decisões, por ficar acordado durante as madrugadas me dando força para que concluísse este trabalho, sobretudo nas últimas semanas que antecederam o depósito da dissertação.

Finalmente agradeço imensamente a Deus e toda a “sua equipe”, que me mantiveram firme e determinado na realização dessa pesquisa.

Resumo

Este trabalho analisa, de acordo com as premissas teórico-metodológicas da Sociolinguística Variacionista (Labov 1972:[2008]), como se dá a concordância nominal de número (CN) na fala de alagoanos estabelecidos na cidade de São Paulo, em comparação com a fala de paulistanos. A análise multivariada é feita com as ocorrências de sintagmas nominais “simples” (constituídos de dois elementos), como “as pessoa-s/ø”. O objetivo principal é verificar quais são os fatores linguísticos e sociais que concorrem para a realização da CN nesses dois subgrupos de falantes. Os dados foram extraídos de 24 entrevistas sociolinguísticas com paulistanos e 24 com alagoanos que vivem na capital paulista. Ambas subamostras são definidas pelas mesmas variáveis sociais: sexo/gênero, faixas etárias e escolaridade.

Os resultados indicam que a taxa de não realização da concordância (CN-ø) entre alagoanos e paulistanos é praticamente a mesma. Em ambas as subamostras, os fatores sociais mostraram-se mais significativos, uma vez que todos foram selecionados para os dois grupos de falantes, com os maiores *ranges*. As mulheres são as mais sensíveis à variante de prestígio; no entanto, as alagoanas apresentam percentual um pouco maior de CN-ø (23%) do que as paulistanas (17%). Não há indicativo de mudança em progresso: em ambos os subgrupos, a segunda faixa etária é a que desfavorece CN-ø. Já para escolaridade, trata-se de um grupo de fatores mais significativo para paulistanos do que para alagoanos. Quanto aos fatores linguísticos, Classe de palavra do elemento nuclear, Processo de formação de plural do elemento nuclear e Número de sílaba do elemento não-nuclear se mostraram mais significativos.

Da perspectiva da concordância nominal de número, esta dissertação mostra que alagoanos e paulistanos se assemelham mais do que se diferenciam. Nesse sentido, pode ser que façam parte de uma mesma comunidade de fala (Labov 1972 [2008]), 1996 [2006], Guy 1981 – mas isso só poderá ser confirmado por estudos futuros, que se dedicarem a outras variáveis na fala destes e de outros grupos de falantes migrantes para a cidade de São Paulo.

Palavras chaves: Concordância Nominal, Variação linguística, Sociolinguística, São Paulo.

Abstract

Within the framework of variationist Sociolinguistics (Labov, 1972 [2008]), this master thesis analyzes variable number agreement within the nominal phrase (NP), in the Portuguese spoken in the city of São Paulo by *Alagoanos* (migrants from the northeastern state of Alagoas) and by *Paulistanos* (those born and raised in the city). The multivariate analysis focuses on two-word plural NPs, such as “*as pessoa-s/ø*” ‘the persons’. The main goal is to verify what linguistic and social factors condition the variable in use, in a comparison between these two groups of speakers. The data was extracted from 24 sociolinguistic interviews with *Alagoanos* and 24 with *Paulistanos*. Both samples are stratified by the same social variables: sex/gender, age group and level of education.

Results indicate that the frequency of CN- $\emptyset$ (that is, lack of agreement) is approximately the same for both groups of speakers. In both, social factors are more significant than linguistic factors, since all of the social ones included in the analysis were selected as statistically significant. Women disfavor CN- $\emptyset$ (a generally stigmatized form), although the agreement is more frequent among *Alagoanos* than among *Paulistanos*. There’s no indication of change in progress: for both groups of speakers, the intermediate age group disfavors CN- $\emptyset$. As for level of education, it is more significant for *Paulistanos* than for *Alagoanos*. Among the linguistic factors, the class of the nuclear word, the morphology of plural, and the number of syllables of the NP left element were the ones selected as statistically significant

This thesis shows that, from the perspective of number agreement within the NP, *Alagoanos* and *Paulistanos* are more similar than different. Therefore, they could be part of a same speech community (Labov 1972 [2008], Guy 1981) in the city – something that should be discussed by further research, including other linguistic variables and groups of speakers.

Keywords: Nominal Phrase, Number Agreement, Language Variation, Sociolinguistics, São Paulo.

Lista de Tabelas

Tabela 1 – Distribuição dos dados dos falantes adultos em função da classe gramatical e da posição dos elementos no SN (transcrita de Scherre, 1996)	20
Tabela 2 – Distribuição das classes gramaticais não-nucleares em função da posição e relação com o núcleo. Scherre (1996:51)	21
Tabela 3: Distribuição geral dos dados alagoanos	64
Tabela 4 Distribuição geral dos dados paulistanos	64
Tabela 5: Análise 1 - CN-Ø entre alagoanos	66
Tabela 6: Análise 1 - CN-Ø entre paulistanos.....	66
Tabela 7: CN-Ø e Faixa Etária	68
Tabela 8: CN-Ø e Faixa Etária para análise atomística e SNs em Oushiro (inédito).....	69
Tabela 9: CN-Ø e escolaridade	70
Tabela 10: CN-Ø e sexo/gênero	72
Tabela 11: Distribuição dos fatores no grupo Formação do plural.....	74
Tabela 12: CN-Ø e formação do plural.....	74
Tabela 13: CN-Ø e classe de palavra do elemento não nuclear.....	76
Tabela 14: Análise 2 – CN-Ø para alagoanos.....	78
Tabela 15: Análise 2 – CN-Ø para paulistanos.....	78
Tabela 16: Análise 3 – CN-Ø para alagoanos.....	79
Tabela 17: Análise 3 – CN-Ø para paulistanos.....	80
Tabela 18: CN-Ø no cruzamento entre Sexo/Gênero e Faixa etária.....	81
Tabela 19 – CN-Ø no Cruzamento entre Sexo/Gênero e Escolaridade.....	83
Tabela 20 – CN-Ø no Cruzamento entre Faixa etária e Escolaridade	85

Lista de Quadros

Quadro 2 : Distribuição dos domicílios por Faixas de Renda Familiar , segundo Distritos do Município de São Paulo (2000) – IBGE – Fundação Seade.....	34
Quadro 3: Grupos de fatores definidos da amostra SP2012 e perfis sociolinguísticos dos informantes.....	40
Quadro 4: Quadro de perfis paulistanos (12.....	41
Quadro 5: Perfis dos informantes alagoanos e paulistanos usados nessa pesquisa	48
Quadro 6: Codificação inicial dos fatores do grupo Classe gramatical do elemento não nuclear	55

Lista de Figuras

Figura 1: Divisão da cidade de São Paulo em Centro Velho, Centro Expandido e Periferia.	33
Figura 2 População adulta (30-60 anos) na Região Metropolitana de São Paulo.....	37
Figura 3: Percentuais de CN-Ø para paulistanos e alagoanos em SNs simples e complexos, e predicativos.....	64
Figura 4: CN-Ø e Faixa Etária para alagoanos e paulistanos (P.R.).....	68
Figura 5: CN-Ø e Faixa Etária em duas pesquisas com amostras paulistanas.....	69
Figura 6: CN-Ø e escolaridade para alagoanos e paulistanos	70
Figura 7: CN-Ø sexo/gênero entre paulistanos e alagoanos.....	72
Figura 8: pesos para CN-Ø de acordo com a formação do plural	75
Figura 9: CN-Ø e classe de palavra do elemento não nuclear	77
Figura 10: CN-Ø e número de sílabas do elemento nuclear (para alagoanos)	81
Figura 11 – CN-Ø no Cruzamento entre Sexo/Gênero e Faixa etária para Alagoanos	82
Figura 12 – CN-Ø no Cruzamento entre Sexo/Gênero e Faixa etária para paulistanos.....	82
Figura 13: Cruzamento entre Sexo/Gênero e Escolaridade para Alagoanos.....	84
Figura 14: Cruzamento entre Sexo/Gênero e Escolaridade para paulistanos	84
Figura 15: CN-Ø no Cruzamento entre Faixa etária e Escolaridade para Alagoanos	85
Figura 16: CN-Ø no Cruzamento entre Faixa etária e Escolaridade para Paulistanos	86

Sumário

AGRADECIMENTOS.....	3
RESUMO	4
ABSTRACT	5
LISTA DE TABELAS	6
LISTA DE QUADROS	7
LISTA DE FIGURAS.....	8
INTRODUÇÃO	10
CAPÍTULO 1 - SOCIOLINGÜÍSTICA VARIACIONISTA E OS ESTUDOS DE CONCORDÂNCIA NOMINAL.....	14
1.1. A teoria da Variação e Mudança.....	14
1.2. A concordância nominal como variável	15
CAPÍTULO 2 – AS AMOSTRAS DE FALA EM SÃO PAULO	30
2.1. São Paulo	32
2.1.1. As imigrações na cidade	36
2.2. O Projeto SP2010	39
CAPÍTULO 3 - A VARIÁVEL E OS GRUPOS DE FATORES.....	49
3.1. Grupos de fatores sociais	51
CAPÍTULO 4. RESULTADOS.....	63
CONCLUSÃO	87
REFERÊNCIAS BIBLIOGRÁFICAS.....	91
ANEXO: ROTEIRO DAS ENTREVISTAS.....	95

Introdução

O português paulistano é uma variedade ainda pouco investigada (Mendes, 2009; Oushiro, 2010). Apesar de significativas empreitadas como o Projeto NURC-SP (Castilho & Preti, 1986, 1987; Preti & Urbano, 1988, 1990), o Projeto Gramática do Português Falado (Castilho, 1990; Castilho *et al.*, 2006; Ilari, 2008) e, mais recentemente, o Projeto SP2010 – Construção de uma amostra da fala paulistana (Mendes, 2012), pode-se dizer, de acordo com Oushiro (2010), que *“há um grande número de trabalhos realizados com base no português paulistano, mas pouquíssimos sobre o português paulistano”*. Em outras palavras: há um grande número de trabalhos que se valem de dados do português paulistano, muitas vezes em estudos comparativos com o português europeu, mas pouquíssimos trabalhos que analisam e descrevem sociolinguisticamente a variedade falada na cidade de São Paulo.

Especificamente sobre o falar paulistano, vem ao caso estudar usos linguísticos (aqui entendidos como “inerentemente variáveis”- (Labov, 1972 [2008])) salientes na comunidade, tais como a pronúncia do /-r/ em coda silábica (Oushiro, 2010), a concordância nominal (Oushiro, inédito) e outras. Entretanto, se por um lado ainda são pouco numerosos os estudos dessa natureza, por outro nem sequer há notícias de trabalhos que incluam a possível convivência, em razão de imigrações, de variedades linguísticas dentro da cidade.

No intuito de contribuir para a ampliação dos estudos sociolinguísticos e a descrição do português falado na capital paulista, este trabalho analisa, de acordo com as premissas teórico-metodológicas da Sociolinguística Variacionista (Labov 1972 [2008]), como se dá a Concordância Nominal (CN) de número em sintagmas nominais de dois elementos (algumas vezes aqui referidos como SNs “simples”, relativamente àqueles que têm três ou mais elementos). As ocorrências de SNs plurais foram extraídas de 24 entrevistas sociolinguísticas com paulistanos e 24 com alagoanos. Nos dados, o primeiro elemento dos SNs é categoricamente marcado com plural, enquanto que a marcação no segundo é variável (como em “*as meninaØ” e “*aqueles moços”*)*

A Gramática Normativa prescreve que todos os elementos do SN plural, assim como um nome e um sintagma predicativo, devem concordar em gênero e número (cf. Bechara, 1999; Cunha, 1971; Cunha & Cintra, 1985; Sacconi, 2004). Entretanto, são inúmeros os trabalhos que já evidenciaram e descreveram o fato de que a concordância nominal de número se realiza variavelmente em toda a comunidade brasileira, verificando-se diferenças relativas a frequências, mas nem sempre com relação aos contextos que favorecem a não realização da concordância (CN-Ø) (Braga, 1977; Scherre, 1988, 1996; Lobato, 1994; Dias, 1996; Scherre e Naro, 1998; Antonino, 2007; Lucchesi, 2008; Salomão, 2010, Fiamengui, 2011 e Oushiro (inédito)).

Alguns resultados têm sido recorrentes em diferentes variedades do português brasileiro em que a variável foi estudada. A posição dos elementos dentro do sintagma nominal (um grupo de fatores especialmente importante na análise de SNs plurais com mais de dois elementos), a classe gramatical e o cruzamento entre esses dois grupos de fatores sempre se mostram relacionados com a variável. Em geral, os determinantes (sobretudo artigos) são os que mais frequentemente carregam a marca de plural (Scherre 1988, Santos 2010, Brandão & Vieira 2012). A Formação do plural (regular, com acréscimo de -s, ou irregular, com -is, ões etc.) também se revela como um grupo de fatores significativo nas análises, podendo variar de comunidade para comunidade (Scherre 1988, Fiamengui 2010) – ainda que, no geral, a não realização da concordância (CN-Ø) seja favorecida nos casos de plural regular.

Além destes e de outros fatores de natureza linguística, a realização variável da CN é sensível a fatores sociais. O nível de escolaridade do informante apresenta-se como um dos fatores que mais favorecem a CN-Ø: quanto menor o nível de escolaridade do informante, maior é a tendência de que ocorra esta variante. A faixa etária dos informantes também tem sido verificada como significativa nos trabalhos, em geral com desfavorecimento de CN-Ø entre falantes da faixa intermediária (34 – 59 anos em média) - embora em algumas comunidades esses resultados possam ser diferentes (Naro & Scherre 1991, Scherre & Naro 2006). Para sexo/gênero – uma das variáveis mais replicadas nos estudos sociolinguísticos – as mulheres tendem a ser as mais conservadoras, enquanto os homens tendem a favorecer CN-Ø, embora nem

sempre este fator seja selecionado como estatisticamente relevante no funcionamento social dessa variável (Salomão, 2010).

Dessa forma, são muitos os trabalhos que mostram como a concordância de número no SN em português brasileiro é um caso de variação socialmente marcada, em que pesam a identidade social do falante (Fiamengui, 2011), em termos de baixos níveis de escolaridade e de informalidade na situação de fala (Scherre, 1988, 1996, Santos, 2012). Em São Paulo, interessa comparar a realização da CN na fala de alagoanos (que migraram para a cidade) e paulistanos (aqueles que nasceram e foram criados nela), a fim de verificar se:

- (i) os fatores sociais e linguísticos que favorecem a ocorrência de CN-Ø são os mesmos nos dois subgrupos;
- (ii) essa variável é estável ou se há indicativo de mudança em progresso (Labov, 1972: [2008]) nos dados extraídos das amostras de fala dos dois grupos;
- (iii) em ambos os subgrupos CN-Ø é de fato a variante favorecida entre as pessoas com menor grau de escolaridade;
- (iv) paulistanos e alagoanos constituem uma mesma comunidade de fala na cidade de São Paulo, no que diz respeito à variável CN.

As 24 entrevistas sociolinguísticas com paulistanos foram selecionadas dentre aquelas gravadas e transcritas pelo Grupo de Estudos Sociolinguísticos da USP (GESOL), na execução do “Projeto SP2010 – Construção de amostra de fala paulistana¹” (Mendes, 2012). Essas entrevistas foram estratificadas de acordo com três variáveis sociais: sexo/gênero (masculino e feminino), faixa etária (20 a 34, 35 a 59 e mais de 60 anos de idade) e escolaridade (Fundamental e Médio). Para cada um dos doze perfis sociolinguísticos que esses fatores geram, o grupo gravou cinco informantes, de modo que o *corpus* totaliza 60 entrevistas, que podem ser acessadas na Internet². Já as 24 entrevistas com alagoanos que residem na cidade de São Paulo foram gravadas e

¹ Proc. FAPESP n°. 2011/09278-6 (pesquisador responsável: Prof. Dr. Ronald Beline Mendes)

² Site: <http://projetosp2010.fflch.usp.br/> (último acesso em 30/07/2014).

transcritas especificamente para o desenvolvimento do presente trabalho, de acordo com os mesmo perfis e orientações gerais do Projeto SP2010.

Esta dissertação está estruturada em quatro capítulos. No primeiro, desenvolve-se uma síntese dos trabalhos realizados por Labov (1972[2008], 1966 [2006]) que fornecem os fundamentos para a pesquisa sociolinguística, bem como dos estudos sobre a concordância nominal que serviram de referência para o desenvolvimento do presente. No capítulo 2, descreve-se como foram construídas as amostras paulistana e alagoana, assim como caracterizam-se os dois grupos sociolinguisticamente. No capítulo 3, são delineadas as hipóteses que se aventaram acerca da realização da CN, sob a forma de grupos de fatores. Finalmente, o capítulo 4 traz os resultados das análises quantitativas dos dados, com foco nos grupos de fatores (sociais e linguísticos) estatisticamente significativos no funcionamento da variável – sempre numa comparação entre alagoanos e paulistanos. Verifica-se que os dois grupos apresentam baixas frequências de CN-Ø, mas que ela é um pouco maior entre paulistanos (afirmação possível graças à comparabilidade entre as duas subamostras). A variação na realização da CN revela-se correlacionada, em ambos os grupos, aos fatores sociais: escolaridade, sexo/gênero e faixa etária. Entre os fatores linguísticos, destacam-se a classe gramatical do elemento à esquerda do SN, a formação (regular ou irregular) do plural no elemento à direita. O número de sílabas do elemento nuclear é um grupo de fatores que foi selecionado apenas entre os alagoanos. No geral, esse estudo comparativo permite dizer que alagoanos e paulistanos mais se assemelham do que se diferenciam, no que diz respeito ao favorecimento de CN-Ø.

Capítulo 1 - Sociolinguística Variacionista e os estudos de concordância nominal

1.1. A teoria da Variação e Mudança

A língua/linguagem como um fato social foi algo proposto pela primeira vez por Saussure (1916 [2000]) em seu curso de linguística Geral. No entanto, o interesse sincrônico pela estrutura linguística, em que o valor da oposição entre unidades se firmou como um dos conceitos centrais contribuiu para que, na prática, a análise de sistemas linguísticos (ou seja, a análise estrutural de línguas diferentes) tivesse lugar privilegiado, relativamente ao estudo de diferenciações intralinguísticas em correlação a fatos de natureza social (ou extralinguística).

Embora a linguística descritiva americana, a partir da década de 20, já apresentasse interesse pela diversidade (intra)linguística dentro de uma comunidade de falantes, os estudos com foco na variação e mudança da língua (do modo como são ainda hoje desenvolvidos) surgem na década de 1960, com Labov, Weinreich & Herzog (1968 [2006]) – que sistematizaram um conjunto de interesses teórico-metodológicos sobre o estudo da mudança linguística com base na variação nos usos. Segundo os autores, a variação está sujeita a pressões internas, emanadas do próprio sistema linguístico, mas também a pressões externas, oriundas do contexto social no qual os falantes estão inseridos. Nas palavras dos autores:

“(...) encontramos razões convincentes para modificar essa posição nos fatos confirmados de que os desvios de um sistema homogêneo não são todos eles erros aleatórios de desempenho, mas são num alto grau codificados e parte de uma descrição realista da competência de um membro de uma comunidade de fala.” (Weinreich, Labov, Herzog, 1968 [2006]:60)

A nova vertente de estudos linguísticos, que se tornaria conhecida como Sociolinguística, passa então a estudar a língua em seu uso efetivo, baseando-se no fato de que a variação é inerente a qualquer manifestação linguística e passível de ser analisada e sistematizada como heterogeneidade estruturada. Weinreich, Labov e

Herzog tinham por objetivo um novo modelo de língua, que acomodasse “os fatos do uso variável e seus determinantes sociais e estilísticos” (1968[2006]:34).

Esse modelo teórico-metodológico de certa forma rompia com as correntes estruturalista e gerativista, na medida em que permitia (senão, pedia) uma nova abordagem para os estudos linguísticos, que se pautasse pelo princípio de que os sistemas são por definição heterogêneos e sujeitos a pressões sociais – de modo que é na heterogeneidade (e não na homogeneidade) que se deve buscar entender a estrutura e o funcionamento da língua.

Dessa forma, desde seus estudos seminais em Martha’s Vineyard (1972 [2008]) e Nova Iorque, Labov (1966 [2006], 1972 [2008]) mostrou que as variantes das variáveis linguísticas são empregadas com significação social e que o funcionamento dos sistemas linguísticos deve ser estudado a partir dessa premissa.

A Sociolinguística laboviana busca, então, descrever e analisar os mecanismos relacionados aos processos de variação e mudança, considerando que os fatores linguísticos (internos à língua) e os sociais (sexo/gênero, escolaridade, faixa etária, nível social, situação de comunicação etc.) podem influenciar a “escolha”³ de uma ou outra variante, mostrando que a língua é regulada por um conjunto de regras variáveis e não categóricas.

1.2. A concordância nominal como variável

De acordo com a prescrição gramatical do português brasileiro (Bechara, 1999; Cunha, 1971; Cunha e Cintra, 1985; Sacconi, 2004), a concordância nominal (CN) é obrigatória, ou seja, se no plural, todos os elementos do sintagma nominal (SN) devem ter a marca de plural, bem como os predicativos a ele ligados. Traduzindo a tradição gramatical, Scherre afirma:

“Na concordância dentro do SN, colocam-se marcas explícitas de plural em todos os seus elementos flexionáveis quando o núcleo do sintagma for formalmente plural; na concordância do predicativo com o sujeito, repetem-se as marcas formais de plural em

³ Não necessariamente consciente.

todos os elementos flexionáveis dos predicativos quando o sujeito for formalmente plural...” (Scherre, 1977:182)

Entretanto, o fato é que, nos usos, essa “marcação” não é categoricamente realizada. Por isso, nos estudos sociolinguísticos, a CN no SN é tratada como uma regra variável. Como tal, a realização da CN é influenciada por fatores sociais e linguísticos. Entre os primeiros, análises variacionistas têm mostrado que a ausência de concordância nominal está comumente relacionada a níveis de escolaridade mais baixos e à informalidade da situação de comunicação. Linguisticamente, destacam-se a posição linear que o elemento ocupa no SN e a classe gramatical de seus elementos.

Braga (1977)

Um dos primeiros trabalhos sobre concordância nominal como variável no PB foi desenvolvido por Braga (1977), que analisou a fala de sete moradores do Triângulo Mineiro (15 a 20 anos de idade) pertencentes às classes média e baixa, com escolarização correspondente aos atuais níveis Fundamental e Médio (incompleto). A autora aventou hipóteses linguísticas que organizou nas seguintes variáveis: (i) *presença e ausência de flexão no elemento anterior ao termo considerado*; (ii) *natureza fonológica do contexto seguinte*; (iii) *posição que o elemento ocupa no SN* e (iv) *grau de saliência fônica entre as formas singular e plural*. As variáveis sociais incluídas na análise foram: *grau de formalidade da entrevista e classe social do falante*.

Os resultados relativos à influência desses fatores na não realização da concordância indicam que são maiores as chances de um elemento do sintagma não receber a marca de plural quando o elemento anterior também não possui tal marca (considerando-se sintagmas nominais com mais de dois elementos)

Sobre a posição que elemento ocupa no SN, Braga (1977) conclui que “os falantes da classe baixa e média apresentam uma probabilidade mais elevada na aplicação da regra (concordância nominal realizada) de presença de concordância nominal na primeira posição, com um declínio progressivo na segunda, terceira, quarta e quinta posições”.

Quanto ao contexto fonológico, Braga (1977) recomenda considerar a natureza fonológica do contexto imediatamente anterior ao /-s/ e a posição da sílaba tônica no elemento considerado-

Já quanto ao grupo *grau de formalidade*, não se observou sua influência entre os falantes da classe baixa, mas sim entre aqueles de classe média. Segundo Braga (1977), os falantes da classe média apresentam desempenho linguístico que se aproxima mais da norma de prestígio do que os falantes da classe social baixa.

Poplack (1980)

Os processos de supressão do morfema plural em elementos do SN ocorrem também no espanhol latino americano. Poplack (1980) analisou 6.349 dados extraídos de 24 entrevistas sociolinguísticas com pessoas da classe trabalhadora acima de 21 anos de idade, falantes do espanhol porto-riquenho e não falantes do inglês no norte da Filadélfia. Suas análises levaram à conclusão de que a presença de marca de plural no segmento anterior favorece a retenção de marca no elemento em questão. Assim como nos resultados de Braga (1977), a noção de pluralidade está associada à cadeia de marcas que existem em elementos precedentes, ou por meio de algum outro mecanismo, como por exemplo, um SN anterior. Assim como mostrou Braga (1977), é significativa a contribuição do fator *posição* para a tendência de supressão da marca do plural no elemento que sucede um outro sem marca. Os exemplos (1) e (2), extraídos das subamostras paulistana e alagoana organizadas para o presente trabalho, ilustram esse fato:

(1) Estar próxima DOS MEUS FILHOS...pra mim...é a melhor coisa desse mundo (SP-F63S)⁴

(2) Ah...era divertido... NAS FESTAØ JUNINAØ a gente fazia fogueira lá no sítio... (AL-M60M)

Segundo as análises de Poplack (1980), um SN do tipo deste sublinhado em (1), com uma sequência em que há marcas na primeira e segunda posição, é bem baixa a taxa de ausência de plural no terceiro elemento. Já em (2), no sintagma tem-se marca

⁴ Nos exemplos, “D1” e “S1” se referem, respectivamente, a Documentador (a) e Sujeito Informante. Os códigos entre parênteses, logo após os exemplos, indicam o perfil do informante: naturalidade do informante (SP – paulistano ou AL – alagoano); sexo/gênero (M – masculino ou F – feminino); a faixa etária (1: entre 20 e 35 anos; 2: 36-59; e 3: 60 anos ou mais); a escolaridade (M para aqueles que estudaram até o Ensino Médio e S para aqueles que estudaram até o ensino superior).

de plural no primeiro constituinte, mas não no segundo - o que aumenta as chances de que também não haja no terceiro. Apenas para o determinante, o primeiro elemento do sintagma, o efeito de supressão é baixíssimo: na maioria das vezes em que não se observa concordância, o primeiro constituinte é o único com a marca /-s/. Neste sentido, Poplack (1980) defende ser esta uma semelhança do espanhol (falado à época na Filadélfia) com o português brasileiro. Este efeito denominado por *paralelismo formal* foi posteriormente testado para o português brasileiro por Scherre (1988).

O estudo de Poplack (1980) fornece ainda outras considerações importantes, relacionadas ao efeito dos fatores gramaticais e dos fatores fonológicos na realização da concordância nominal. No que diz respeito às classes de palavras, constata que na posição não nuclear à esquerda do SN, um adjetivo favorece a não realização da concordância, enquanto que um artigo a favorece. Já com relação aos fatores fonológicos, constata que uma pausa seguinte ao SN plural favorece CN-Ø, do mesmo modo que tonicidade fraca na sílaba seguinte.

Pode-se considerar que a conclusão mais relevante do trabalho feito por Poplack (1980) é a de que, enquanto o espanhol da Filadélfia (elimina a redundância da estrutura superficial da sentença), ela organiza o seu sistema de marcação de plural. Na variedade falada, o espanhol analisado parece necessitar apenas que o plural seja marcado em algum constituinte do sintagma, diferentemente do que prescreve a gramática normativa.

Na presente pesquisa, com base na fala paulistana e alagoana, o grupo de fatores *posição* e *marcas precedentes* não serão testadas. Ao serem extraídas todas as ocorrências de sintagma nominal das duas amostras, constatou-se que a quantidade de SNs simples (dois elementos) ocorrem em média cinco vezes mais do que os SNs mais complexos (três elementos ou mais). Como o interesse maior dessa pesquisa reside na comparação entre os dois grupos de falantes, optou-se por não analisar os SNs com mais de dois elementos. Por outro lado, Oushiro (inédito), que também trabalha com essa variável numa amostra paulistana mais abrangente, apresenta análises que incluem também os SN complexos). No presente trabalho, verifica-se que, nos SNs simples, a primeira posição é categoricamente marcada com plural, diferentemente da segunda. Embora não se trabalhe com os dois grupos de fatores acima mencionados, faz-se necessário apresentar resultados já obtidos sobre ele, pois

além de se mostrar muito significativo para a CN-Ø, está diretamente relacionado com o grupo de fator *Classe de palavras do elemento não nuclear*, que se inclui nas análises da variável entre alagoanos e paulistanos.

Scherre (1988 e 1996)

Scherre (1988) analisa o efeito da variável *paralelismo formal* e verifica que a presença de marca de plural num determinado elemento do SN é favorecida pela presença de tal marca nos elementos anteriores. No mesmo sentido, a presença de pelo menos uma marca precedente não sucedida por zero (ou seja, ausência de marca de plural) favorece a presença de marca no elemento seguinte, com percentual um pouco menor do que quando há duas ou três marcas antecedendo o elemento em questão. Já a presença de zero num elemento precedente leva quase categoricamente ao zero no constituinte seguinte.

Scherre (1996) discorda da correlação sugerida por Guy (1981) - de que há uma equivalência entre os fatores classe de palavra e posição do elemento no SN - e desenvolve uma abordagem que considera três grupos de fatores linguísticos (*posição linear, classe gramatical, marcas precedentes*). A partir de 11.086 dados, extraídos de 48 entrevistas sociolinguísticas do *Corpus Senso*⁵, procura verificar todas as inter-relações entre eles e os transforma em dois fatores: (1) *relação entre elementos nucleares e não nucleares e posição linear dos elementos nucleares do SN* e (2) *marcas precedentes em função da posição*.

Com o objetivo de explicar para o efeito do fator classe gramatical (Guy, 1981) em função do fator posição, Scherre (1996) apresenta uma tabela com a distribuição dos dados, a fim de embasar sua discordância com Guy (1981):

⁵ Corpus Censo do Programa de Estudos sobre o Uso da Língua (PEUL), vinculado ao Departamento de Linguística e Filologia da Faculdade de Letras da Universidade Federal do Rio de Janeiro - UFRJ. Esse banco de dados é constituído por 64 gravações de 60 minutos cada com informantes estratificados em função dos anos de escolarização (de 1 a 4 anos, de 5 a 8 anos; e de 9 a 11 anos), da faixa etária (7/14 anos; 15/26; 26/49 e acima de 49 anos), e do sexo (feminino e masculino). (cf. SCHERRE, 1996).

Tabela 1 – Distribuição dos dados dos falantes adultos em função da classe gramatical e da posição dos elementos no SN (transcrita de Scherre, 1996)

Classe	Posição			Total
	Primeira	Segunda	Terceira	
Substantivos	153/161 = 95% <i>Ex.: <u>Problemas</u> assim maiores</i>	2.777/5.196 = 53% <i>Ex.: Essas <u>carne</u> congelada</i>	317/514 = 62% <i>Ex.: das otas <u>famílias</u></i>	3247/5871
Categoria substantivada	Não ocorre	60/93% = 65% <i>Ex.: trabalha uma, as <u>outras</u> fica</i>	16/23 = 70% <i>Ex. as coisas <u>boas</u> e as ruins</i>	76/116
Pronome pessoal 3ª pessoa	12/12 = 100% <i>Ex. <u>delas</u> todas</i>	27/33 = 82% <i>Ex. de todos <u>eles</u></i>	Xxxx <i>Ex. todos <u>ele</u></i>	39/45
Adjetivo	47/48 = 98% <i>Ex.: <u>novas</u> escolas</i>	125/163 = 77% <i>Ex.: as <u>boas</u> ações</i>	125/307 = 41% <i>Ex: uma coisinha <u>bonitinha</u></i>	297/518
Quantificador	133/133 = 100% <i>Ex.: <u>todos</u> os anos</i>	3/7 = 43% <i>Ex.: eles <u>todo</u></i>	22/108 = 20% <i>Ex.: aquelas cruzinha <u>toda</u></i>	158/248
Possessivo	184/184 = 100% <i>Ex.: <u>suas</u> tias</i>	130/135 = 96% <i>Ex.: nos <u>meus</u> filhos</i>	2/12 = 17% <i>Ex.: uns colega <u>meu</u></i>	316/331
Demais classes não nucleares	3701/3802 = 97%	141/151 = 93%	3/4 = 75%	3845/3953
Total	4230/4340 = 97%	3263/5778 = 56%	485/968 = 50%	7978/11086

A autora mostra que a terceira posição dos SNs não é ocupada apenas por adjetivos (31% são adjetivos e 53% são substantivos), e que 30% dos adjetivos aparecem na segunda posição, percentual significativamente alto para ser ignorado; os substantivos têm maior probabilidade de serem marcados na terceira posição (62%), ao contrário dos adjetivos que se apresentam mais marcados na segunda (77%); e, por fim, os possessivos e demais elementos não nucleares também ocorrem na segunda posição com consideráveis índices de marcas (97% e 93%). Além disso, verifica-se ainda que qualquer que seja a classe gramatical que ocupe a primeira posição, esta tende a ser muito marcada.

Logo, a autora verifica que tomar classe por posição ou posição por classe leva a encobrir regularidades linguísticas importantes. A partir dessa argumentação, ela

introduz uma nova abordagem analítica que contempla a relação entre essas duas variáveis, uma vez que: (i) as classes gramaticais não nucleares que ocorrem na primeira posição são todas antepostas ao núcleo do SN; (ii) as classes gramaticais da segunda posição são predominantemente antepostas ao núcleo do SN; e (iii) as classes não nucleares da terceira posição são predominantemente pospostas ao núcleo do SN.

Para essa análise, a autora apresenta uma tabela com as classes não nucleares distribuídas em função da posição linear que ocupam no SN em função da relação de anteposição ou posição de núcleo do SN.

Tabela 2 – Distribuição das classes gramaticais não-nucleares em função da posição e relação com o núcleo. Scherre (1996:51)

CLASSE E POSIÇÃO	Anteposto	Posposto
Adjetivo na 1ª posição (<i>novas escolas</i>)	47/47 = 98%	Não ocorre
Adjetivo na 2ª posição (<i>as boas ações/coisas lindas</i>)	36/39 = 92%	89/124 = 72%
Adjetivo na 3ª posição (<i>uma casinha bonitinha</i>)	Não ocorre	127/307 = 40%
Quantificador na 1ª posição (<i>todos os anos</i>)	133/133 = 100%	Não ocorre
Quantificar na 2ª posição (<i>eles todo</i>)	Não ocorre	3/7 = 43%
Quantificador na 3ª posição (<i>aquelas cruzinha toda</i>)	Não ocorre	22/108 = 20%
Possessivo na 1ª posição (<i>suas tias</i>)	184/184 = 100%	Não ocorre
Possessivo na 2ª posição (<i>os meus filhos/colegas minhas</i>)	127/130 = 98%	3/7 = 43%
Possessivo na 3ª pessoa (<i>uns colegas meu</i>)	Não ocorre	2/12 = 17%
Demais elementos não-nucleares 1ª posição (<i>alguns meses/determinados pontos</i>)	3791/3802 = 97%	Não ocorre
Demais elementos não-nucleares 2ª posição (<i>todos os anos/eles mesmo</i>)	141/148 = 95%	0/3 = 0%
Demais elementos não-nucleares 3ª posição (<i>do meus próprios filhos</i>)	1/1 = 100%	2/3 = 66%

De acordo com seus resultados, relativamente às classes não nucleares foi possível observar que todos os elementos que antecedem o núcleo do SN são marcados e que todos os elementos *pospostos* ao núcleo do SN são menos marcados.

Scherre (1996) considera ainda que a relação entre classe gramatical e posição, relativamente aos elementos não nucleares, deve ser vista por meio de sua distribuição em torno do núcleo, sem que importe tanto nem a classe nem a posição, mas sim a distribuição da classe não nuclear em relação ao centro do SN.

Segundo a autora, revela-se importante uma segunda questão: a posição linear mostra-se significativa somente com relação aos elementos não nucleares do SN; tal constatação é interessante na medida em que: as categorias substantivadas comportam-se da mesma forma que os substantivos, sendo mais marcados na terceira do que na segunda posição; e pronomes pessoais evidenciam maior número de marcas na primeira do que na segunda posição, embora apresentem porcentagem mais alta do que os substantivos. A autora propõe, então, unir os dois grupos de fatores em apenas um, que denomina *Relação entre os elementos do SN*:

1 – as classes não nucleares antepostas apresentam mais marcas do que as pospostas; não é o adjetivo que é pouco marcado, mas sim o adjetivo posposto; não é o quantificador que é muito marcado, mas o quantificador anteposto, justificando-se que não é correto afirmar que “sempre” a primeira posição do SN é mais marcada que as demais.

2 – Os elementos não nucleares não são igualmente marcados em todas as posições, já que primeira e a terceira posição são sempre mais marcadas do que a segunda, refutando-se conclusões de que a classe dos substantivos é uma das menos marcadas.

Scherre (1996) retoma os fatores *marcas precedentes* e *posição linear*, estudados tanto no PB quanto no espanhol, com o objetivo de mostrar que ambas as línguas apresentam mais semelhanças do que diferenças. Ao analisar esses fatores para o PB, a autora constata que não foram analisados adequadamente, indicando que os autores não discutiram os resultados obtidos. Ela dividiu *marcas precedentes* em nove fatores e os exemplos por ela apresentados estão destacados abaixo, nos quais o elemento em foco encontra-se sublinhado e o *contexto precedente* apresenta-se grafado com letras maiúsculas:

- o fator *ausência* refere-se à ausência de qualquer constituinte do SN precedendo o elemento da primeira posição (*os fregueses, do meus tio*).

- zero formal na primeira posição – (*dO meus tios; FILHOTINHO novos; TANTA gargalhadas*).
- numerais da primeira posição - (*TRÊS capítulos; NOVE núcleos formados; DEZ senhoras lá sentada; DOIS risco verde*).
- presença de marca formal na primeira posição - (*OS fregueses; UNS troço; OS meus quatro filhos; AS perna toda marcada; NOTAS melhores*).
- núcleo semântico e presença de marca formal – engloba casos como: (a) de núcleo nominal mais alto com marca formal de plural (*MILHÕES de coisas*); e (b) os casos de núcleo nominal, com ou sem marca formal de plural, acompanhado de marca no primeiro elemento do SN em análise (*UM GRUPO DE CRIANÇAS abandonadas; UMA SÉRIE DE OUTRAS coisas*).
- núcleo semântico e ausência de marca formal – apresenta (a) as ocorrências de núcleo nominal mais alto sem marca formal de plural (*UMA PORÇÃO DE carro; UM GRUPO DE crianças abandonadas*); e (b) as ocorrências de núcleo nominal sem marca formal de plural seguido de ausência de marca no primeiro elemento do SN sob análise (*UMA PORÇÃO DE COISA interessante*).
- presença de marcas formais a partir da primeira posição — envolve a presença de duas ou mais marcas formais de plural precedendo o segmento analisado na terceira, quarta ou quinta posição (*AS MAIORES privações; AS PARTIDAS TODAS iguais*).
- mistura de marcas – engloba todos os casos de presença de pelo menos uma marca formal de plural precedendo o elemento nominal analisado, não mediada por zero em elemento que admite marca (*dO MEUS tios; NOVE NÚCLEOS formados*), os demais casos deste fator (*DEZ SINHORAS LÁ sentada; AQUELAS PESSOAS ASSIM BEM esquisitinha*) seguem o mesmo princípio.
- zero formal a partir da primeira posição – pode ser caracterizado nos seguintes termos: entre a última marca formal e o elemento analisado tem de haver, necessariamente, um zero em elemento que admite marca, mesmo que mediado por um numeral ou por um modificador (*UMAS BORRACHA grande; DOIS RISCO verde; AS PERNA TODA marcada; AS CASA MAIS antiga*).

As análises com esses fatores, com exceção de *ausência fora do SN* e *zero na primeira posição*, que respectivamente favorecem marcação do plural na primeira e na segunda posição, levam à constatação recorrente de que: “marcas levam a marcas e

zeros levam a zeros” (Scherre, 1989). O núcleo semântico formalmente marcado (*MILHARES DE coisas*) ou seguido de marca formal (*UMA SÉRIE DE OUTRAS coisas* e *UMA PORÇÃO DE CRIANÇAS abandonadas*) favorece a inserção de marca de plural no segmento analisado (peso .74); com o núcleo semântico formalmente sem marca de plural (*UMA PORÇÃO DE coisa*) ou seguido de um elemento também sem tal marca (*UMA PORÇÃO DE PEIXINHO colorido*) favorece o zero no elemento seguinte (.41). Dessa forma, segundo Scherre (1989), não são necessárias explicações *ad hoc*.

A autora finalmente conclui que o fator *posição* isolado dos fatores *marcas precedentes* e da *classe gramatical* não dá conta de explicar a variação na concordância de número, na sua totalidade. Reitera, assim, sua hipótese de que a melhor análise para esse fenômeno é a que considera esses três fatores agrupados em dois: 1) *marcas precedentes em função da posição*; e 2) *relação entre elementos nucleares e não nucleares e posição dos elementos nucleares dentro do SN*.

Carvalho (1997)

Carvalho (1997) analisou a concordância de número no SN (simples e complexos) no português falado na área urbana de Rio Branco (Acre). Além de verificar o fenômeno como variável, buscou demonstrar se os contextos linguísticos e sociais correlacionados a ela são os mesmos para aquela comunidade, se comparados com os trabalhos sobre CN no Brasil. Para isso, a autora testou os grupos de fatores: *posição do elemento do SN, classe gramatical, marcas precedentes, contexto fônico seguinte, saliência fônica (tonicidade e número de sílabas), sexo, grau de escolarização e grau de formalidade do discurso*. Sua amostra é composta de 24 informantes, sendo 12 homens e 12 mulheres, com idade entre 20 e 30 anos.

Seus resultados mostram que, para *saliência fônica*, os itens mais salientes são mais marcados com plural do que os menos salientes. Quanto à tonicidade, a autora conclui que não se trata de um grupo de fatores explanatório, já que seus resultados apresentam discrepância entre os percentuais e os pesos relativos (segundo ela, devido à má distribuição dos dados). A hipótese acerca do *número de sílabas do item lexical singular* (itens lexicais com mais de duas sílabas têm mais chance de receber

marca de plural do que os dissílabos e monossílabos) confirma-se. A autora verifica que o princípio da saliência fônica atua mais fortemente sobre a concordância do que *processos morfológicos de formação de plural*.

Carvalho (1997) afirma, assim como Scherre (1988, 1996), que as variáveis *posição, classe gramatical e marcas precedentes* apresentam-se muito inter-relacionadas; assim *Classe gramatical*, relacionado à *posição dos elementos no SN*, demonstra que os determinantes são mais marcados nas duas posições e a comparação entre substantivos e adjetivos revela que os primeiros são mais marcados com plural do que os segundos, nas duas posições.

Com relação às variáveis sociais, a autora verificou que o *grau de escolarização*, está intrinsicamente ligado ao fenômeno variável, numa correlação que se repete nos trabalhos - quanto maior o nível de escolaridade do informante, menor é a tendência de não realização da concordância. Quanto ao sexo, os dados não confirmam a hipótese de que as mulheres mais sensíveis ao prestígio, uma vez que são os homens que menos deixam de realizar a concordância. A autora afirma que tal resultado pode estar relacionado ao engajamento dos homens de sua amostra no mercado profissional.

Andrade (2003)

Entre os numerosos trabalhos sobre variação na concordância nominal, há também aqueles desenvolvidos na região sul do Brasil, tais como Andrade (2003), que analisou 24 entrevistas sociolinguísticas gravadas em Tubarão, no Sul de Santa Catarina, e em São Borja, no Rio Grande do Sul (12 informantes de cada cidade). As entrevistas de São Borja fazem parte do Projeto VARSUL⁶ e as de Tubarão foram cedidas pelo Procotextos/Amurel⁷.

Foram consideradas as variáveis *idade* (entre 20 e 60), *sexo* (masculino e feminino) e *escolaridade* (primário, ginásio e colegial). Linguisticamente, os grupos foram: *posição do elemento no SN, Classe gramatical dos elementos do SN, relação*

⁶ Projeto VARSUL - Variação Linguística Urbana da Região Sul (UFSC).

⁷ PROCOTEXTOS/AMUREL - Projeto de Coleta de Texto de Informantes da AMUREL.

desses elementos com o núcleo, marcas precedentes, processo morfológico de formação de plural, tonicidade dos elementos e graus dos substantivos e adjetivos.

O cruzamento entre *posição linear* e *classe gramatical*, assim com nos trabalhos apresentados até aqui, possibilitou verificar que a primeira posição do SN é um fator que favorece a inserção da marca de plural, o que já não acontece com as demais posições, que tendem a desfavorecer a presença de marcas de plural no SN. No entanto, Andrade (2003:107) evidencia uma diferença em relação a trabalhos anteriores — *“quando a segunda posição é ocupada por artigos ou demonstrativos, esta irá favorecer mais a aplicação da regra que a primeira posição ocupada por esta mesma classe gramatical”*. Ou seja, na primeira posição, os substantivos favorecem mais a aplicação da regra do que os próprios determinantes; já os adjetivos desfavorecem a aplicação da regra em qualquer posição.

Assim como em outros trabalhos, plurais irregulares favorecem mais a aplicação da regra (isto é, a realização da concordância) do que os plurais regulares. Quanto à variável *marcas precedentes*, a autora confirma a hipótese de Scherre (1988) de que *marcas levam a marcas e zeros levam a zeros*. Os resultados do grupo de fatores *relação com o núcleo* comprova que todos os elementos antepostos ao núcleo do SN são muito mais marcados, enquanto os elementos pospostos são pouco marcados.

Em relação à classe gramatical, quando substantivos aparecem na primeira posição do SN, apresentam peso relativo maior para a não realização da concordância - maior que os determinantes na primeira posição que a desfavorecem.

Dentre as variáveis sociais, *escolaridade* levou a resultados semelhantes ao de trabalhos anteriores. As mulheres deixam de realizar a concordância relativamente menos que os homens - na cidade de Tubarão (SC), mas não em São Borja (RS), onde não se verifica diferenciação entre homens e mulheres. A idade dos informantes não se revelou um grupo de fatores relevante.

Fiamengui (2011)

A variação na CN em SNs também foi estudada no interior paulista. Aqui, destaca-se o trabalho de Fiamengui (2011), com foco na fala de adolescentes da região de São José do Rio Preto. Foram analisados 5.391 dados extraídos de duas amostras:

640 redações escolares obtidas a partir do projeto “Desenvolvimento de Oficinas de Leitura e Interpretação de Textos no Ensino Fundamental”⁸ e 8 narrativas orais integrantes do corpus Iboruna, constituído no âmbito do Projeto ALIP⁹.

Nesse estudo, foram considerados os seguintes fatores linguísticos: *posição linear e classe gramatical, processo morfológico de formação de plural, tonicidade do item lexical singular, número de sílabas do item lexical singular, contexto fonético/fonológico seguinte, marcas precedentes ao elemento analisado, formalidade dos substantivos e adjetivos*.

Já os fatores sociais foram: *sexo (masculino e feminino), idade (10 a 15 anos), escolaridade (5ª a 8ª séries do ensino fundamental)*. Para a análise das redações, foi considerado ainda o grupo *proposta de redação* e para a comparação entre as duas amostras incluiu-se o grupo *modalidade (oral/escrita)*.

Os resultados de Fiamengui (2011) mostram que a modalidade escrita desfavorece a não realização da concordância, relativamente à oral. Considerando-se aqui apenas os dados orais, o grupo *posição linear e classe gramatical* apresentou-se como o mais significativo, sendo que o determinante (anteposto, na primeira posição do SN) é aquele em que a marca de plural é mais presente, seguido pelo determinante anteposto ao núcleo ocupando a segunda posição. Os adjetivos pospostos só favorecem a marcação do plural quando são trissílabos a partir da terceira posição do SN.

O grupo de fatores *número de sílabas do item lexical* não confirma a hipótese de que os itens com maior número de sílabas são mais marcados com plural que os demais. A autora considera que uma possível explicação para esse resultado possa estar na posição que o item ocupa no SN. O cruzamento entre os grupos *posição linear e classe gramatical e número de sílabas* mostra que a maior parte dos monossílabos

⁸ Projeto desenvolvido na Escola Estadual Professora Zulmira da Silva Sales, localizada na área residencial e periférica da cidade de São José do Rio Preto, credenciado junto à Pró-Reitoria de Extensão da UNESP, da mesma cidade, e coordenado pelas Prof^{as} Dra. Luciani Ester e Prof^a Dra. Sanderléia Roberta Longhin-Thomazi.

⁹ ALIP – Amostra Linguística do Interior Paulista – Banco de Dados construído entre 2004 e 2007 com a amostra da fala de 152 informantes da região de São José do Rio Preto, estratificados em sexo/gênero (masculino e feminino), faixa etária (07 a 15 anos, 16 a 25 anos, 26 a 35 anos, 36 a 54 anos e mais de 55 anos), escolaridade (1º ciclo do ensino fundamental, 2º ciclo do ensino fundamental, ensino médio e ensino superior) e renda (mais de 25 sal. min., 11 a 24 sal. min., 6 a 10 sal.min. e até 5 sal. min.)

ocupa a primeira posição (substantivos ou categorias substantivadas), caso no qual a marcação do plural é desfavorecida.

O último grupo selecionado entre os linguísticos foi *marcas precedentes ao elemento analisado*. Os resultados permitem, mais uma vez, verificar a atuação do paralelismo formal (marcas levam a marcas e zeros levam a zeros).

Dentre os grupos sociais, *sexo/gênero* foi o mais significativo, sendo que os homens são mais conservadores, ou seja, tendem a deixar de realizar a concordância menos que as mulheres. Fiamengui justifica esse resultado pelo fato de que a maior quantidade de dados é oferecida por informantes de 14 anos, que na amostra influenciam negativamente a marcação.

Os resultados para *idade* são interessantes, uma vez que os mais velhos entre seus informantes (14 anos) desfavorecem a marcação de pluralidade. Essa é justamente a idade em que, segundo Labov (1974), o indivíduo passa a dispor de maior “percepção social”; segundo Fiamengui (2011:129), “*um dos motivos dessa recusa pode estar numa relação conflituosa entre os valores incalculados pelo sistema escolar com a adoção da norma prescritiva nas aulas e língua portuguesa, e os valores culturais que o adolescente herda de sua própria comunidade, dentre os quais os relacionados ao próprio vernáculo*”.

Oushiro (inédito)

Na capital paulista, Oushiro (inédito) realiza uma pesquisa com dados de fala paulistana, em uma amostra composta por 118 entrevistas sociolinguísticas, com os mesmos perfis dos paulistanos aqui analisados, em que analisa, entre outras variáveis, a concordância nominal de número. A autora analisa tanto sintagmas simples como sintagmas complexos, com os seguintes fatores: *sexo/gênero, faixa etária, escolaridade, região de residência, classe social, origem dos pais, mobilidade geográfica, processos morfofonológicos de formação do plural, tonicidade do item singular, posição relativa, classe morfológica, contexto fônico seguinte e paralelismo intrassintagmático (somente para 3ª, 4ª e 5ª posições)*.

Dentre as variáveis sociais, não são selecionadas como significativas para a concordância nominal a faixa etária do falante, a sua região de residência e a origem dos pais. A Idade do falante, como variável contínua, tampouco é selecionada. Esses resultados indicam que, no quadro geral, a concordância nominal parece ser um caso de variação estável em São Paulo.

Dentre as variáveis linguísticas *posição linear e classe de palavras* foram mais significativas. Quanto ao paralelismo sintático – investigado por Scherre & Naro (1992), Scherre (1998) e Scherre (2001) – os resultados são similares: SNs precedidos por outros SNs com concordância parcial favorecem CN- $\emptyset$ (peso .70), e SNs precedidos por SNs com concordância plena a desfavorecem (peso 45); quando se trata do primeiro SN de uma série (neste estudo, a ausência de SNs plurais nas cinco orações precedentes), a tendência é neutra, semelhante à probabilidade geral da amostra.

A amostra analisada por Oushiro (inédito) é muito maior e mais abrangente do que aquela que se analisa aqui (pensando-se nos paulistanos). Entretanto, o objetivo do presente trabalho é, conforme vem-se reiterando, comparar a realização da CN na fala de dois grupos de falantes que vivem na cidade de São Paulo. Nesse sentido, ambos os dois trabalhos cooperam na descrição da fala paulistana. A revisão bibliográfica apresentada neste primeiro capítulo fornece as bases para a formulação de hipóteses, na forma de grupos de fatores – muitos dos quais são incluídos na análise comparativa que aqui se desenvolve.

Capítulo 2 – As amostras de fala em São Paulo

Desde seus trabalhos seminais, Labov (2008 [1972]) ressalta que a realização de estudos de variáveis linguísticas em áreas urbanas não se pode basear em dados fornecidos por informantes selecionados arbitrariamente. Além de uma descrição da população que se quer focar, no sentido de compreender ao menos algumas particularidades da comunidade, deve-se cuidar para que haja, de acordo com os interesses do estudo, variação também na amostra da comunidade: homens e mulheres, de idades diferentes, pertencentes a estratos socioeconômicos distintos, que desenvolvam atividades ocupacionais variadas. Naturalmente, a hipótese de base é que essas diferenças podem atuar nos usos linguísticos dos membros de uma tal comunidade.

De acordo com o autor, é preciso compreender quais as fronteiras que delimitam o enquadramento de um indivíduo a uma certa comunidade de fala, e não a outra. Considera que as fronteiras são postas mediante dois aspectos: um deles, no nível consciente, e o outro, no inconsciente. Quanto ao nível consciente, os falantes compartilham atitudes e valores semelhantes em relação à língua, já que a comunidade de fala é mais bem definida como um grupo que compartilha as mesmas normas em relação à língua (LABOV, 1972 [2008]). Tais normas são apreendidas pelo pesquisador mediante o “valor” que os falantes de uma certa comunidade de fala atribuem a elas, sendo que, normalmente, ao grupo de prestígio – cuja fala é dominante na escola, no trabalho, na mídia, etc. – são vinculados valores positivos.

Ainda de acordo com Labov (1972), além de valores conscientes em relação à língua, os falantes de uma comunidade de fala compartilham, inconscientemente, aspectos essenciais do sistema linguístico – as suas regras gramaticais –, sendo que os indivíduos adquirem tal sistema sem que possam “escolher” falar deste ou daquele jeito.

Ao considerar esses dois níveis envolvidos na delimitação da comunidade de fala, Labov ressalta que ela não pode ser simplesmente concebida como um grupo de

falantes que utiliza as mesmas formas, mas como um grupo que compartilha as mesmas normas em relação à língua (1972:158). Dessa forma, é possível observar que ele prioriza o caráter de consciência das atitudes dos falantes em relação às normas gramaticais compartilhadas pelo grupo, para caracterizar a sua comunidade de fala.

No presente trabalho, alagoanos e paulistanos em São Paulo podem constituir comunidades de fala distintas (com hábitos sociais e linguísticos particulares). Entretanto, a verificação disto dependeria não só da inclusão, no debate, de outros grupos de migrantes que vivem na cidade (baianos, pernambucanos, etc), além de incluir também outras variáveis linguísticas. De fato, não é tarefa trivial discutir o que define uma comunidade de fala, sobretudo num espaço urbano tão heterogêneo quanto São Paulo. Neste sentido, a análise que se desenvolve aqui é uma contribuição na direção de desenhar um quadro mais abrangente da sociolinguística paulistana, que passe a incluir aqueles que se mudaram para a cidade, além dos paulistanos nativos. Em outras palavras, verificar se paulistanos e alagoanos, na cidade de São Paulo, se assemelham ou se diferenciam quanto ao emprego da concordância nominal é mais um passo em tal direção.

Para tanto, construiu-se uma amostra de 24 entrevistas sociolinguísticas gravadas com alagoanos que moram em São Paulo e selecionaram-se 24 entrevistas com paulistanos da amostra coletada pelo GESOL (SP2010) . Conforme descrevem as seções a seguir, essas duas amostras (que por vezes são referidas aqui como subamostras, quando se considera o conjunto total de dados para o desenvolvimento desta dissertação) são estratificadas com base nas variáveis sexo/gênero (masculino e feminino), faixa etária (20 a 34, 35 a 59 e mais de 60 anos) e dois níveis de escolaridade (Ensino Fundamental e Ensino Médio). Desses fatores sociais, resultam 12 perfis sociolinguísticos diferentes (por exemplo: homem de primeira faixa etária com ensino médio), para cada um dos quais há dois informantes.

Labov (1972[2008]) sugere que uma amostra de falantes idealmente conta com pelo menos três indivíduos para cada perfil. Entretanto, o trabalho de localizar, contactar e gravar certos perfis de alagoanos (por exemplo, aqueles de terceira faixa etária que tenham concluído o ensino fundamental) não permite que se colete uma amostra tão grande num tempo limitado. Dessa maneira, gravaram-se dois informantes por perfil.

Conforme indica o Capítulo 1, a novidade do presente trabalho está na comparação entre dois subgrupos de falantes que habitam São Paulo. Como os nativos da cidade são foco também de outra pesquisa – que também enfoca a concordância nominal de número (além de outras variáveis) e que utiliza uma amostra muito maior e mais representativa (118 informantes, estratificados também de acordo com a região da cidade) – destaca-se aqui a construção de uma amostra da fala de alagoanos que se mudaram para São Paulo. A próxima seção neste capítulo, contudo, trata da amostra paulistana, já que foi nos conhecimentos acumulados pelo GESOL, na construção da amostra SP2010, que a coleta dos dados alagoanos se baseou.

A seguir, portanto, após uma breve descrição sociolinguística da cidade de São Paulo, este capítulo trata da construção da amostra de alagoanos em São Paulo e faz uma revisão das hipóteses que justificam os grupos de fatores sociais que estratificam ambas as subamostras.

2.1. São Paulo

Situada na região sudeste do país, São Paulo foi fundada em 25 de janeiro de 1554, pelos padres José de Anchieta e Manuel de Nóbrega, que haviam subido a Serra do Mar em busca de um local seguro para se instalar e catequizar os índios. Quase cinco séculos depois, o povoado de Piratininga transformou-se na maior metrópole brasileira, com 1500 km², sendo hoje a maior cidade da América do Sul, com uma população de 11,8 milhões de habitantes. É também a principal cidade da Região Metropolitana do estado de São Paulo, rodeada por outras, que formam a Grande São Paulo (aproximadamente 19,7 milhões de habitantes).¹⁰

A capital paulista é hoje o grande polo econômico do país; sua economia é baseada na prestação de serviços, seguido pela indústria; já a agricultura não é desenvolvida na cidade, mas sim em outras no interior estado, que abastecem a capital.

¹⁰ Fonte: <http://www.cidades.ibge.gov.br/xtras/perfil.php?lang=&codmun=355030&search=sao-paulo/sao-paulo> – acessado em 26/07/2014

A cidade é dividida em cinco zonas (Centro, Sul, Norte, Leste e Oeste) e estas se dividem em diversos distritos. Para a amostra SP2010, construída pelo GESOL e disponibilizada na Internet, os perfis sociolinguísticos têm, cada um, cinco informantes - um de cada uma dessas zonas. Além dessa distribuição, os informantes da amostra também se dividem em três regiões: Centro Velho, Centro Expandido e Periferia. O *Centro Velho* é o centro histórico (e inclui bairros considerados “tradicionais”, tais como Bela Vista, Consolação, Bom Retiro, Brás e Liberdade) – cuja urbanização é antiga, cujas construções têm um caráter histórico e cujas famílias que aí vivem tendem a ser daquelas mais prototipicamente paulistanas (Mendes; Oushiro, 2012). O *Centro Expandido* também faz parte de uma região central (isto é, não periférica), mas já não constitui o “centro” propriamente. Ela inclui bairros como Pinheiros, Santana, Vila Madalena, Moema e outros, que são relativamente mais desenvolvidos, concentram imóveis que são mais valorizados e contam com melhor infraestrutura de serviço e cultura. Já a *Região Periférica* (que inclui bairros como Sapoemba, Pirituba, Jabaquara, Campo Limpo e Perus) é aquela cuja formação é historicamente mais recente; assim, os bairros não dispõem da melhor infraestrutura, comparativamente às outras duas regiões. A figura abaixo ilustra essa divisão dos bairros da cidade:

Figura 1: Divisão da cidade de São Paulo em Centro Velho, Centro Expandido e Periferia.
(Mendes & Oushiro 2012)

São Paulo apresenta distribuição socioeconômica bastante contrastante, já que é comum encontrar numa mesma região distritos altamente valorizados em que se concentram as maiores rendas familiares (como no Morumbi, mais de 60% habitantes detêm mais de 25% salários mínimos por família) (Mendes & Oushiro, 2012) e distritos menos valorizados (como em Parelheiros, em que 64% detêm renda familiar até 5 salários mínimos)¹¹. O quadro abaixo mostra a distribuição de renda familiar da população em diferentes distritos da cidade. É comum que em distritos mais afastados das regiões centrais os índices de renda familiar sejam bem menores do que em distritos centrais:

Distritos/Região (C/P)	Faixas de Renda (em Salários Mínimos – SM)					
	Até 2 SM	2-5 SM	5-10 SM	10-15 SM	15-25 SM	Mais de 25 SM
MUNICÍPIO DE SÃO PAULO	13,30	24,39	25,97	11,29	10,98	14,06
Capão Redondo (P)	18,30	32,80	29,82	9,36	6,50	3,22
Consolação (C)	3,77	8,33	17,19	10,49	15,74	44,47
Grajaú (P)	22,92	35,87	28,21	7,55	4,08	1,38
Itaim Bibi (C)	3,87	6,61	12,98	8,83	14,78	52,92
Jardim Paulista (C)	2,43	4,04	11,97	9,17	16,61	55,78
Marsilac (P)	44,10	31,23	19,47	4,31	0,19	0,71
Parelheiros (P)	26,06	38,16	25,79	6,15	2,83	1,01

Quadro 1 : Distribuição dos domicílios por Faixas de Renda Familiar , segundo Distritos do Município de São Paulo (2013) – IBGE – Fundação Seade

Segundo os próprios paulistanos que fazem parte da amostra SP2010, as diferentes regiões da cidade podem apresentar características particulares. Há informantes que acreditam que os hábitos das pessoas, inclusive os linguísticos, são diferentes entre as zonas centrais e periféricas, o que pode ser uma consideração

¹¹ Fonte: <http://www.cidades.ibge.gov.br/xtras/perfil.php?lang=&codmun=355030&search=sao-paulo|sao-paulo> – último acesso em 26/07/2014

importante para a investigação da concordância nominal, já que se trata de uma variável estigmatizada e que pode traçar perfil social dos falantes.

(3) **D:** *you consegue nota(r) diferença de comportamento... alguma coisa... entre os paulistanos de lugares diferentes da cidade?*

S: *sim...com certeza dá dá pra...é o paulistano ele...eu não sei se o paulistano...o ser humano em si...ele é muito assim pré-moldável moldável assim pra muda(r) inclusive a sua característica dependendo de onde ele está ele se comporta de uma forma diferente...porém tendo escapes que ele vê que ele não aguenta mais e ele acaba mostrando quem ele é realmente mas tem mudança de comportamento mediante a cada lugar que vai sim*

D: *é (vo)cê acha que a gente acaba se moldando*

S: *acaba com certeza eu acho que é do próprio ser humano isso né tem locais que você deve... que você tem que manter uma postura diferenciada como uma pessoa que mora num bairro i(r) procura(r) emprego no centro da cidade ele precisa ele vai muda(r) o seu comportamento a sua forma de fala(r) porque senão ele não vai consegui(r) emprego né e mesmo porque eu acredito que tem que existi(r) profissionais pra pode(r) faze(r) avaliação né e é de fato dinâmicas de grupo entrevistas não tem essa necessidade (SP-F2M)*

(4) **D:** *e o que que (vo)cê acha que tem de diferente no jeito de fala(r) de alguém do centro ou de diferentes bairros de periferia mesmo?*

S: *o vocabulário bastante diferenciado o vocabulário não as gírias né as gírias você vê que é diferente ahn no meu ver é a gíria né são gírias diferentes né são vocabulários regionais*

D: *mesmo dentro da cidade tem também?*

S: *tem tem sim com certeza (SP-M2S)*

Para o interesse da pesquisa sociolinguística na cidade de São Paulo é importante ressaltar que, embora “Classe social” não seja um grupo de fatores que participe da estratificação das amostras, pode estar intrinsicamente relacionada a Escolaridade, já que quanto maior o nível de escolaridade do informante maiores são as probabilidades de acesso a melhores posições profissionais e redes sociais. Embora a amostra paulistana tenha incluído variabilidade em termos de classe social, este grupo de fatores não será considerado no presente trabalho análise pela dificuldade de estratificar a amostra alagoana da mesma forma e de encontrar alagoanos de diferentes classes sociais na cidade.¹²

¹² Cf. Oushiro (inédito) para uma análise da relevância do grupo de fatores Classe no funcionamento da concordância nominal em São Paulo.

2.1.1. As imigrações na cidade

De acordo com recente levantamento do Instituto de Pesquisa Econômica Aplicada (IPEA, 2013), 46% da população residente da região metropolitana não nasceu na cidade – algo que fica bem ilustrado na Figura 1, abaixo. As migrações de estrangeiros para São Paulo foram impulsionadas sobretudo durante o período da Segunda Guerra Mundial, que interrompeu as importações e levou a um aumento da produção, no estado, dos produtos até então importados. O processo intensificou-se no governo de Juscelino Kubitschek, que lançou as bases da indústria automotiva no ABC paulista, um tema que também aparece em entrevistas com paulistanos, gravadas pelo GESOL:

(5) *D: e essa questão da imigração... mais... né... tinha essa questão da Liberdade bairro de japonês... o Brás e a Móoca bairro de italiano... (vo)cê acha que ainda (es)tá concentrado assim... (xxx)... nos bairros... ou já misturou?*

S: não.. tem muita mistura... mas tem muita concentração porque eu... trabalhei na Secretaria da Cultura né...é então... é... tem muita... imagina... o pessoal... o povo judeu é super fechado... o pessoal da... da Itália... eles são também super organizados...

D: os judeus (es)tão aonde?... assim... tem um bairro?...

S: os judeus?... olha... acho que Higienópolis tem muito... tem...eles... que(r) ve... lá perto da Hebraica também... é... ele não tem assim um lugar acho que o judeu como tem por exemplo a Liberdade ou como o Bexiga...o italiano... mas eles são muito unidos sabe... eles... é uma comunidade que é bem... muito forte... e os (japoneses) também... não tem aquelas festas de vez em quando que tem...fazem aquelas festas como que chama... da Achiropita né?... né... é... mas tem... tem... tem as fes... tem uma festa que... é lá no... como é que chama... aquele museu... Memorial do Imigrante... uma vez por ano... tem a Festa das Nações... lá você vê as comunidades né... e eles dançam... todo ano tem... é maravilhoso...eles fazem stands com comidas típicas... e as comunidades elas preparam shows típicos pra apresenta(r) no palco... então é... tem muita mistura... mas tem muita organização nas comunidades independentes... (en)tendeu?... (SP-M1S)

Conforme já se mencionou neste trabalho, o português falado em São Paulo foi pouco estudado, e a cidade é hoje constituída por pessoas de diversas nacionalidades e naturalidades. Os imigrantes mais enraizados na cidade (japoneses, italianos, coreanos, bolivianos) em bairros como Mooca, Liberdade, Bom Retiro, não tiveram ainda o seu português estudado. Embora não trabalhe diretamente com eles, Oushiro (inérito) investiga a concordância nominal e outras variáveis na fala de uma variedade abrangente de paulistanos, que possivelmente inclui descendentes de italianos - o que

abre caminhos para trabalhos futuros. A fala dos migrantes nordestinos que residem na cidade de São Paulo também não foi estudada ainda, e os dados do gráfico abaixo revelam que grande parte da população paulista é constituída por nordestino. A presente pesquisa, evidentemente, busca contribuir para ampliar o cenário dos estudos sociolinguísticos na cidade.

Figura 2 População adulta (30-60 anos) na Região Metropolitana de São Paulo.

Fonte: Censo 2011 (*apud* Rocha, 2013)

De acordo com esse gráfico, cerca de 30% da população residente na Região Metropolitana de São Paulo são representados por nordestinos. De acordo com os dados disponibilizados pelo Centro de Tradições Nordestinas (CTN), o total de nordestinos que reside na cidade é de 2.047.168¹³. Para suprir a necessidade de mão de obra operária, o estado passou a receber migrantes nordestinos a partir da metade do século XX, em especial homens, vindos principalmente dos estados da Bahia, Ceará, Pernambuco e Paraíba, que se fixaram principalmente na periferia de São Paulo e nas cidades vizinhas. No processo de gravação das entrevistas, tanto da amostra alagoana quanto da paulistana, é comum que os informantes identifiquem alguns bairros ou regiões da cidade em que se concentram mais imigrantes ou mais nordestinos, justamente por estes serem mais numerosos na cidade.

(6) *D: aqui é verdade... você acha que tem alguma região assim que concentra.... sei lá os nordestinos pessoas pessoas...do nordeste?*

¹³ disponível em <http://www.ctn.org.br> – último acesso em 20.08.2013

S1: é no... Santo Amaro...Santo Amaro... se concentra a maior região de nordestino lá você se sente que (vo)cê (es)tá no nordeste...tem até lugares assim que vende coisas típicas do nordeste lá...

D: ah é?

S: tem salão de forró... tem tudo... do nordeste...é um pouquinho longe né... ali Largo Treze de Maio...se bem que hoje em dia está difícil de dizer que nordestino se concentra só em um lugar...antes o centro era só pra pessoas...pessoas que tinham mais condições...hoje...se você andar pela região do Brás...por exemplo...(vo)cê vai encontrar muito nordestino...principalmente porque lá tem muita agência de ônibus clandestinos... (AL-M2M)

No período entre 1970 a 1991, mais de 2,6 milhões de nordestinos migraram para a cidade de São Paulo, com maior número de baianos, seguidos de pernambucanos, cearenses, paraibanos e alagoanos que viam na cidade a possibilidade de melhores condições de vida:

“Dentre os imigrantes registrados em São Paulo nas décadas de 70 e 80, constata-se que, no primeiro período, a número de imigrantes da UF chegava a cerca de 3 milhões, representando 92% do total de imigrantes do Estado, principalmente devido aos volumes de emigrantes da região Sul do país. Já no período seguinte os imigrantes não nascidos do Estado foram cerca de 2,3 milhões sendo que houve um pequeno acréscimo no número de imigrantes de retorno que chega a 368 mil pessoas, ou seja, cerca de 14% do total.” (Cunha et al.,2000: 71-72)

Os dados sobre a cidade de São Paulo, sobretudo sobre as migrações, reforçam o carácter multicultural da cidade e sua complexidade etnográfica. Não é possível estudar aspectos socioculturais e linguísticos da cidade sem levar em consideração o convívio dos diferentes subconjuntos de habitantes. Por este motivo, neste trabalho serão analisadas amostras da fala de dois desses subconjuntos que convivem na capital paulista: paulistanos nativos (aqueles que nasceram e foram criados na cidade) e alagoanos que se mudaram para ela. Conforme adianta a introdução, estuda-se como se realiza a concordância nominal de número nessas duas amostras, a fim de verificar se a variação é regulada pelos mesmos grupos de fatores sociais e linguísticos.

O item a seguir trata dos procedimentos para a construção da amostra de 60 entrevistas por meio do projeto “SP2010 – Construção de uma mostra da fala paulistana” (Mendes e Oushiro, 2012), das quais foram selecionadas 24 para a composição da amostra paulistana utilizada nesta pesquisa.

2.2. O Projeto SP2010

Por volta de 1970, iniciaram-se no Brasil os primeiros estudos que investigavam a língua falada no Brasil. O Projeto NURC (Norma Linguística e Urbana Culta) (Castilho & Preti, 1986, 1987; Preti & Urbano, 1988, 1990) possibilitou que diversos estudos linguísticos fossem realizados com algumas das variedades do português falado no país, incluindo-se entre elas a falada em São Paulo.

No estado de São Paulo, alguns trabalhos que disponibilizam banco de dados de fala foram realizados e merecem destaque, como Projeto para o Português Paulista (Castilho, 2007) e o Projeto Iboruna (Gonçalves, 2008), na região de São José do Rio Preto.

Entre os anos de 2008 e 2010, o GESOL construiu uma amostra da fala paulistana composta por mais 100 entrevistas com paulistanos e não paulistanos, de ambos os sexos, entre 15 e 89 anos de idade e com escolaridade equivalente ao Ensino Fundamental incompleto até o Ensino Superior completo, residentes das cinco zonas da cidade (Centro, Norte, Sul, Leste e Oeste). Tal amostra é chamada de “Piloto” e o período de sua coleta possibilitou que roteiros de entrevistas fossem reelaborados e aprimorados, métodos de contato de informantes fosse refinados e critérios de transcrição e de codificação de variáveis fossem revistos.

A partir de análises dessas entrevistas da Amostra Piloto, consideraram-se seis os parâmetros mais importantes para a construção de uma amostra representativa da fala paulistana Sexo/Gênero, Faixa Etária, Escolaridade, Classe Social, Região da Cidade e Geração da Família na cidade. No entanto, não seria factível construir uma amostra que se estratificasse com base em todas essas variáveis em um período limitado. Para manter certa comparabilidade com outras amostras construídas em outras localidades, decidiu-se que a nova amostra (SP2010) seria estratificada por: sexo/gênero (masculino e feminino); faixa etária (20-34 anos; 35-59; 60 anos ou mais) e escolaridade (Ensino Fundamental e Ensino Médio). Essas variáveis levam a 12 perfis, para cada um dos quais foram gravados 5 informantes, totalizando-se 60 entrevistas, do modo como resume o quadro 02. Quanto as outras três variáveis não incluídas na

estratificação dos perfis, elas não definem o tamanho da amostra, mas fazem parte dela qualitativamente.

Quadro 2: Variáveis estratificadoras da amostra SP2012 e perfis sociolinguísticos dos informantes (transcrito de Mendes e Oushiro 2012: 978).

Sexo/Gênero (F/M)	Faixa Etária (1/2/3)	Escolaridade (C/S)	Perfil sociolinguístico
Feminino	20-34 anos	Até Ens. Médio	1 . F1C
		Ens. Superior	2 . F1S
	35-59 anos	Até Ens. Médio	3 . F2C
		Ens. Superior	4 . F2S
	60+ anos	Até Ens. Médio	5 . F3C
		Ens. Superior	6 . F3S
Masculino	20-34 anos	Até Ens. Médio	7 . M1C
		Ens. Superior	8 . M1S
	35-59 anos	Até Ens. Médio	9 . M2C
		Ens. Superior	10 . M2S
	60+ anos	Até Ens. Médio	11 . M3C
		Ens. Superior	12 . M3S
12 perfis x 5 informantes = 60 gravações			

Conforme já se adiantou na Introdução, essas 60 entrevistas estão disponíveis (tanto o áudio quanto as transcrições) em site próprio¹⁴ na Internet. Dentre elas, para o presente trabalho foram selecionadas 24 entrevistas, com base nos perfis e informantes que seriam gravados para a amostra alagoana – que é descrita no item a seguir. Em princípio, a proposta para a construção da amostra alagoana era idêntica à paulistana; no entanto, não foi possível encontrar alagoanos da terceira faixa etária com ensino superior completo, de modo que alterou-se a variável escolaridade (em vez de Ensino Médio e Superior, trabalhou-se com Ensino Fundamental e Médio). Tal alteração não representou um problema para a amostra paulistana, que contém grande variedade de perfis e informantes (considerando-se não apenas a Amostra SP2010, mas também a Amostra Piloto).

¹⁴ Cf: <http://projetosp2010.fflch.usp.br/>

Sexo/Gênero	Faixa etária	Escolaridade	Perfis
Feminino	1 (20-35)	Fundamental	F1M
			F2M
			F3M
	2 (36-59)	Médio	F1S
			F2S
			F3S
Masculino	1 (20-35)	Fundamental	M1M
			M2M
			M3M
	2 (36-59)	Médio	M1S
			M2S
			M3S

Quadro 3: Amostra Paulistana (12 perfis x 2 informantes = 24)

Todos os informantes da amostra SP2010, conforme já se mencionou, nasceram e foram criados na cidade de São Paulo. Já vimos também que a metrópole é habitada por uma grande quantidade de imigrantes, que deveriam passar a integrar os estudos sociolinguísticos. Neste sentido, o presente trabalho dá um passo, ao coletar uma amostra de alagoanos que vivem na capital São Paulo e ao propor a comparação de suas falas, ainda que da perspectiva de uma única variável linguística – a concordância nominal de número.

2.3. A construção de uma amostra de alagoanos em São Paulo

A comunidade alagoana na cidade de São Paulo soma 145.237 indivíduos, (66.760 homens e 78.447 mulheres¹⁵). Assim como é comum entre nordestinos migrantes, eles vieram para a metrópole em busca de melhores ofertas de trabalho e melhores condições de vida para seus familiares. Com o passar do tempo, foram se enraizando e construindo relações sociais e afetivas com a cidade de São Paulo – o que fez com que sentissem acolhidos, mas nem por isso deixassem de preservar costumes, conforme exemplificam os trechos de entrevistas abaixo:

¹⁵ De acordo com dados disponíveis em: <http://www.ctn.org.br> – último acesso: 20.05.2014.

(7) **D:** *you like to live here in São Paulo why?*

S: *eu gosto porque São Paulo tem tudo e ninguém se mete na sua vida... eu acho isso uma das coisas mais maravilhosas da Terra... você tem seus direitos seus deveres sem interferência das pessoas porque assim...eu estou comparando com o Nordeste...no Nordeste as cidades são pequenas... as pessoas todas se conhecem... qualquer ato que você faz que para você não seja inflacionário... aos olhos das pessoas... seja você entra na mira das pessoas e as pessoas te matam por antecendência...São Paulo não... São Paulo não é todo mundo. (ALM2M)*

(8) **D:** *and you had or maintain some tradition... anything that you did in family?*

S: *tínhamos de...da páscoa né? que eles chamam é a páscoa aquela coisa da gente almoçar na casa de um... almoçar na casa de outro... do afilhado levar um presente para o... para a madrinha e ganhar um presente também né?... eu lembro que na época da...da páscoa era uma época muito gostosa para essa tradição de uma família almoçar na casa da outra... ou cear que seria a janta né? e era bem gostoso eu lembro que era uma dess/ uma tradição que a família tinha e até hoje ainda tem é essa lá... eles mantêm essa tradição aqui também. (ALM1M)*

Na busca da preservação de costumes, conforme menciona a informante ALM1M e de acordo com conversas com integrantes do Centro de Tradições Nordestinas (CTN), os nordestinos, embora tenham se habituado ao modo de vida paulistano, não deixaram de preservar e praticar as suas tradições e costumes na cidade para a qual se mudaram. Em geral, dispõem de centros culturais, restaurantes, igrejas e festas populares, através dos quais vão se mantendo as suas raízes, assim como revela o informante abaixo:

(9) **D:** *is there any place here in São Paulo where the alagoanos gather to preserve their customs?*

S: *então para os nordestinos eu conheço...conheço não é conheço não né... sei que tem o CTN...centro e tradições nordestinas...estive lá uma vez assim mesmo por que vieram parentes minhas do nordeste e a gente foi lá mas faz pouco tempo tenho vontade de ir lá...sempre falo...mas...mas nunca vou né dizem que eles tem uma comida muito boa é e fora isso eu sei que tem pessoas de vários lugares...festas...danças...igreja. (ALM2M)*

Situado à Rua Jacofer, no bairro do Limão, o CTN é amplamente conhecido e reconhecido, tanto pelo público que lota suas instalações todo final de semana (cerca de cem mil frequentadores por mês), como pelo poder público: o centro foi declarado como Entidade de Utilidade Pública Municipal, pela prefeitura de São Paulo, e como Organização da Sociedade Civil de Interesse Público (OSCIP), pelo Governo Federal.

Além disso, a população nordestina é marcadamente conhecida pela religiosidade, conforme indica o trecho da entrevista transcrito abaixo:

(10) D: e vocês tinham alguma tradição religiosa?

S: bom.... meus pai sempre tiveram muito tradição religiosa de família... né não/... sempre foi uma família muito religiosa... lá em Alagoas é... a crença mais é a religi/ religião católica né... o catolicismo...o povo vai a missa....tem as procissão pela cidade....as festas dos boiaderos....vão a Juazeiro na igreja de Padim Ciço....(AL-M1M)

São Paulo dispõe hoje de um local especialmente construído para os nordestinos professarem sua fé – A igreja da Imaculada Conceição, localizada no bairro do Limão, fundada em agosto de 1997 em homenagem a Frei Damião, com capacidade para acomodar 400 fiéis. Nessa igreja existem duas estátuas do venerado Padre Cícero, o “Padim Ciço”, verdadeiro mito nordestino. Está aí também exposto um carro Ford, modelo 1929, usado pelo padre para suas peregrinações pelo sertão nordestino. Esses itens todos, de certa forma, simbolizam o modo como não se transportaram para São Paulo apenas as pessoas, como se tudo mais houvesse ficado para trás, mas foram trazidos também elementos marcantes de sua identidade (em termos amplos). Esses elementos, assim como os falares das pessoas, contribuem para a formação de uma nova realidade na metrópole – tanto do ponto de vista daqueles que se mudaram para ela, quanto do ponto de vista da própria São Paulo.

Monitoradas pela gerente de cultura Mariana Conti¹⁶, as visitas realizadas ao CTN, permitiram conhecer melhor os alagoanos que residem em São Paulo. Por ser um local onde os oriundos de Alagoas e de outros estados nordestinos sentem-se à vontade para manifestar sua cultura, seus costumes e sua linguagem, sem a preocupação com a possibilidade de avaliação negativa por parte de outrem, algumas das entrevistas foram gravadas ali. Pôde-se ter contato direto com informantes de diferentes faixas etárias. Observou-se que, embora a grande maioria dos migrantes tenha vindo a São Paulo em busca de melhores empregos, na verdade este objetivo nem sempre é alcançado (de acordo com dados do CTN), devido às poucas oportunidades de estudo que tiveram na cidade natal durante a infância. Em seus

¹⁶ A quem agradeço pela hospitalidade e disposição em fornecer informações sobre a cultura nordestina e ajudar no contato com os informantes alagoanos.

depoimentos, os entrevistados dizem que a vida na cidade natal não possibilitou que pudessem concluir os estudos e por isso desempenhavam serviços rurais ou domésticos. Também alegam ter poucas condições de concorrer a melhores vagas de emprego tanto na cidade de origem quanto em São Paulo. Muitos encontraram na metrópole uma oportunidade de concluir os estudos básico e superior, o que possibilitou uma melhor colocação profissional para alguns:

(11) **D:** *Você concluiu o ensino médio?*

S: *na realidade eu fiz... é... no sítio eu estudei praticamente uns três ano... de três a quatro ano só que... esses três quatro ano que eu fiquei no sítio... bem dizer não valeu nada porque eu não aprendi nada... né... aí... é... repetia de ano direto...é... aí quando eu cheguei aqui em São Paulo... é... eu cheguei a concluir com supletivo...né...eu cheguei a concluir com supletivo (ALM1M)*

(12) **D:** *Até qual série você estudou?*

S: *lá eu fiz até o magistério é lá né? quando eu vim para cá para São Paulo eu já tinha a formação de magistério então quando eu comecei a trabalhar aqui na área de educação comecei com o magistério né? depois que eu fiz a pedagogia né? mas foi o magistério que...que me colocou na área de educação (ALM2S)*

Ainda que a vinda para a cidade de São Paulo tenha facilitado a conclusão do ensino médio ou superior para alguns, foi difícil encontrar, durante o processo de coleta das entrevistas, informantes com a estratificação estabelecida pelo Projeto SP2010. De fato, entre os informantes contactados, foi bastante reduzido o número daqueles que tivessem ensino superior concluído, sobretudo entre aqueles com mais de 60 anos de idade.

A gravação das entrevistas (tanto dentro como fora do CTN) foi viabilizada graças ao método da observação participante (Malinowski, 1976) e à manifestação do interesse em desenvolver pesquisa sobre os alagoanos residentes na cidade. Isso facilitou significativamente o contato com alguns informantes. Alguns deles foram muito acolhedores e chegaram a oferecer suas casas para a realização das entrevistas. Em sua grande maioria, residem nas regiões periféricas da cidade de São Paulo, com especial destaque para as zonas Sul (Campo Limpo, Macedônia, Parelheiros), Leste (Artur Alvin, Penha, Itaquera) e Norte (Cachoeirinha e Eliza Maria) – o que evidencia que pontos da cidade com maior infraestrutura não são normalmente ocupados por nordestinos. Os entrevistados informaram que a grande maioria dos seus familiares

mora em regiões metropolitanas (Franco da Rocha, Taboão da Serra, São Bernardo, Embu das Artes, Mauá), cidades próximas da capital onde o custo de vida e de moradia são mais baixos se comparados ao das regiões centrais.

Nas entrevistas, os alagoanos manifestam que o convívio social diário com paulistanos modificou e modifica gradativamente o seu modo nordestino de se comportar e de falar. Em geral, mencionam que não tiveram oportunidade de estudar e que, “por isso não conseguem falar o português dos paulistanos”, que tendem a avaliar como “mais correto”. Nesse sentido, os grupos de fatores (sociais e linguísticos), dos quais se trata a seguir, podem permitir verificar se, pelo menos no que diz respeito à concordância nominal de número, quão diferentes (ou semelhantes) são os paulistanos e alagoanos. O contato com os alagoanos durante as visitas ao CTN permitiu verificar que um “modo mais nordestino” de falar (sintaxe, dialetos e sotaques) emerge em conversas com amigos, parentes e conterrâneos, ao passo que, na situação de entrevista propriamente, tendem a aproximar sua fala do (que entendem ser) “padrão”. De um ponto de vista metodológico, essa variabilidade (esperada e atestada em muitas pesquisas sociolinguísticas) também aparece ao longo de uma mesma entrevista, já que há momentos, por exemplo, em que um mesmo informante está mais pessoal e emotivamente conectado a um determinado assunto e, daí, afasta-se do considerado padrão, conforme deixa de monitorar sua própria fala.

As entrevistas com alagoanos em São Paulo foram gravadas e transcritas entre 2012 e 2013 e seguiu os mesmos critérios estabelecidos pelo GESOL para a gravação das entrevistas do Projeto “SP2010”. Os informantes alagoanos foram selecionados de modo semialeatório, a partir do modelo “amigo de amigo” (Milroy, 2004) em que o contato com o informante pode ser intermediado por um terceiro. Dessa forma pode-se estabelecer uma interação com o sujeito informante que pode, após a gravação de sua entrevista, indicar algum amigo ou conhecido que pode também se disponibilizar a ceder uma entrevista sociolinguística. Também foram utilizadas as “redes sociais” (Facebook e Orkut) e visitas ao Centro de Tradições Nordestinas (CTN) para se contactar informantes, já que não é tão comum dentro de uma cidade tão grande como São Paulo encontrar, com facilidade, alagoanos dispostos a conversar por cerca de uma hora com uma pessoa desconhecida. Esses métodos de localização de

informantes garantem alguma aleatoriedade e abrangência, uma vez que se passa a conhecer uma variedade de pessoas que pertencem a grupos sociais diversificados.

Após o primeiro contato com o potencial informante, agenda-se uma conversa, esclarecendo a ele/ela que sua fala será gravada e transcrita, com o objetivo de registrar o vernáculo (como se estivesse conversando com familiares ou em ambientes descontraídos), que posteriormente é analisado em um trabalho acadêmico. Um grande entrave para se conseguir registrar uma fala “menos monitorada” está principalmente no fato de que não há familiaridade entre entrevistador e entrevistado. Por este motivo, foi necessário que o primeiro se colocasse como integrante da comunidade – como um filho de alagoanos, que muitas vezes visitou várias cidades do estado. Isso contribuiu para que as conversas fossem menos “tensas” e possivelmente menos automonitoradas do que poderiam ser e, portanto, mais próximas do vernacular.

O roteiro de entrevista utilizado pelo GESOL precisou ser adaptado de forma a possibilitar que os informantes alagoanos, embora estivessem falando de São Paulo, pudessem discorrer sobre a sua cidade natal, seus costumes, família, atividades de natureza econômica realizadas tanto em São Paulo quanto em Alagoas, suas características diferenciais dos paulistanos e dos outros nordestinos que se encontram na cidade. Informações como essas foram importantes para se desenhar uma etnografia da comunidade alagoana na cidade, e para se levantar hipóteses sobre a correlação entre usos linguísticos e variedades sociais diversas. O roteiro de entrevista do GESOL divide-se em três partes (cf. anexo I). A primeira trata de assuntos gerais que incluem:

- O bairro em que o informante mora ou morou por mais tempo;
- Infância e amizades;
- Família;
- Vida profissional e educação;
- Atividades de lazer.

Para esta primeira parte (que é a mais longa), foi necessário fazer algumas adaptações ou fazer um paralelo entre São Paulo e Alagoas, pois há alguns informantes que estão aqui desde crianças e há aqueles que vieram já depois de adultos. Logicamente, a visão sobre determinado ponto da “discussão” é diferente a depender da extensão de tempo na cidade e da relação de afetividade com ela.

A segunda parte busca apreender algumas percepções e avaliações do falante sobre a cidade de São Paulo, em especial sobre como eles veem a cidade. Já a terceira faz algumas incursões nas visões linguísticas que o falante tem sobre a sua própria fala e sobre a fala do outro (pessoas que não são naturais da cidade), com questões do tipo: *“como é que as pessoas falam na cidade de São Paulo? (cf. Anexo I)*

Ao final dessa terceira parte, pede-se que o informante faça a leitura de dois textos (um deles com características orais) e de uma lista de palavras. As ocorrências de SNs plurais nestes não seriam computadas, mas ainda assim mantiveram-se as leituras nas entrevistas, para o interesse de trabalhos futuros – sobretudo acerca de variáveis fonológicas.

Inicialmente, projetou-se uma amostra da fala de alagoanos em São Paulo com 36 indivíduos (3 por perfil sociolinguístico). Entretanto, a dificuldade para localizar informantes alagoanos nos perfis comparáveis aos paulistanos gravados pelo GESOL, bem como o tempo que tal tarefa tomou, fizeram com que a amostra se limitasse a dois informantes por perfil. O tempo para a localização, contato e gravação dos informantes se alargou para além do previsto principalmente porque a maioria dos alagoanos, conforme vimos, está distribuída pela região metropolitana, concentrando-se nas periferias. Além disso, foi difícil encontrar alagoanos de segunda e de terceira faixa etária com Ensino Superior completo, de forma que foi necessário adaptar a variável escolaridade de forma a garantir pelo menos um informante de cada nível. Colocando alagoanos ao lado de paulistanos num quadro que resume as duas subamostras, temos o seguinte:

Perfil	Paulistanas	Alagoanas	Perfil	Paulistanos	Alagoanos
F1F		1	M1F		
F1M	3	2	M1M	3	3
F1S	1	1	M1S	1	1
F2F	1	1	M2F	1	1
F2M	2	2	M2M	1	1
F2S	1	1	M2S	2	2
F3F	2	3	M3F	2	2
F3M	2	1	M3M	2	2
F3S			M3S		

Quadro 4: Número de alagoanos e paulistanos por perfil

Como é possível verificar, não se encontraram informantes mais velhos de nível superior de escolaridade, de modo que, para fins de comparabilidade, não se selecionou nenhum informante com perfil semelhante da amostra paulistana (evidentemente, eles existem na amostra de sessenta informantes do Projeto SP2010, conforme é possível verificar no site em que se disponibiliza o *corpus*). Por outro lado, gravou-se uma alagoana com nível fundamental, mas, no Projeto SP2010, o nível de escolaridade mais baixo entre as paulistanas é médio. Esses fatos acerca da escolarização dos informantes são retomados no momento em que se apresentam resultados das análises quantitativas.

Paralelamente às gravações e transcrições das entrevistas com alagoanos, fez-se a extração de ocorrências de sintagmas nominais plurais em 24 entrevistas com paulistanos, disponibilizadas pelo GESOL-USP, com perfis o mais semelhantes possível dos perfis alagoanos, a fim de se ter uma amostra balanceada. A tabela a seguir apresenta os 48 perfis (alagoanos e paulistanos) e a quantidade de informantes por células utilizados na composição da amostra completa.

Capítulo 3 - A variável e os Grupos de Fatores

Segundo Labov (2008 [1972]), os objetivos de um estudo sociolinguístico são descrever e analisar estatisticamente um fenômeno linguístico variável, do modo como amostrado numa determinada comunidade de fala. Uma variável é constituída por diferentes modos de dizer uma mesma coisa, por parte dos falantes de tal comunidade. Conforme vimos nas primeiras páginas dessa dissertação, seu interesse é a realização variável da concordância nominal de número (CN *versus* CN-Ø) na fala de alagoanos que moram em São Paulo, em comparação à sua realização na fala dos naturais da cidade - os paulistanos. O interesse dessa comparação, também conforme vimos anteriormente, está no fato de que, por um lado, estudos têm mostrado que são numerosos, através do território brasileiro, os casos em que os mesmos grupos de fatores são selecionados como significativos, com hierarquias semelhantes de correlações dos fatores com a variável dependente. Por outro lado, Oushiro (inédito) discute que a não realização da CN pode ter um significado social local em São Paulo (considerando-se, entre outros fatos, que seus informantes tendem a descrever a variante CN-Ø como algo “paulistano”). Assim, comparar a realização da CN na fala de alagoanos em SP e paulistanos, além de permitir verificar se os mesmos grupos de fatores (tanto sociais quanto linguísticos) se correlacionam à variável, também contribui com a ampliação e compreensão do quadro de trabalhos sobre CN no Brasil.

A partir da leitura das transcrições e da audição das 24 entrevistas com alagoanos e 24 com paulistanos, foram extraídos todos os SNs (com dois constituintes ou mais) e também as construções predicativas. Estas últimas e os SNs com mais de dois elementos são apresentados no capítulo 4 (da perspectiva das distribuições de dados de CN-Ø), mas não são estatisticamente analisados. Conforme já se delimitou na introdução, as análises quantitativas propriamente se restringem aos SNs de dois elementos, como nos exemplos abaixo:

(13) “...minha sogra tem OUTRAS CASAS de aluguel AS CONDIÇÕES não permitiam e a gente acabou ficando aqui mesmo” (AL-F1M)

(14) “aqui já tinha... ALGUMAS RUAS...quando viemos mora(r) aqui já tinham algumas sim...” (AL-F1M).

(15) “...porque eu não queria mora(r) aqui né...ESSAS PESSOAØ ai no (XXX)...e eu queria volta(r) para lá mas ai...mas nós fomos acostumando” (SP-F2M)

(16) “aí encontrei um/ ela...ela não é professora ela cuida assim DAS CRIANÇAØ na hora do intervalo... e ela chama S. também.” (SP-F2M).

Alguns dados potenciais não foram incluídos na análise, como nos casos exemplificados a seguir:

i. SNs compostos por palavras originalmente estrangeiras (a fim de evitar a possibilidade da influência do sistema de concordância dessas línguas no fato de que a concordância foi ou não realizada – já que a maioria desses termos são empréstimos do inglês):

(17) ah...sempre tinha UNS GAMES pra gente joga(r) (SPM2M)

(18) é um espaço para acontecer essas...ESSES SHOW no é quando é época de política (ALF2M)

(19) come(r) no KFC... no Burger... essas coisas... aí você vai NUNS SHOPPINGS vai até... (SPM1M)

ii. SNs cujo elemento nuclear é um numeral (que não admite flexão de plural):

(20) “quando voltei era por volta DAS CINCO...cinco e meia da tarde” (ALF2F)

(21) “então...ai eu dei um para cada um DOS DOIS e fui embora” (SPM3M)

iii. SNs cujo elemento nuclear é seguido por palavras cujo segmento fônico inicial não permite determinar se foi ou não pronunciado um /-s/de plural:

(22) “ESSAS CENAS geralmente não saem da minha cabeça” (ALF3F)

(23) “...prefiro acreditar que AS PESSOAS são boas até que provem o contrário (ALM3M)

Também não foram considerados os SNs cuja pronúncia em volume muito baixo ou em contexto ruidoso impossibilitou determinar com clareza se a concordância foi realizada ou não.

O capítulo 1 apresenta uma breve revisão de alguns dos muitos estudos que analisaram a realização variável da concordância nominal de número no PB. Nesse conjunto de trabalhos, verifica-se a influência de alguns fatores como favorecedores ou desfavorecedores de CN-Ø. Com base neles, mas sobretudo em Scherre (1988, 1994, 1996), a análise comparativa dos dados extraídos de entrevistas com alagoanos e paulistanos inclui os seguintes grupos de fatores: *(i) sexo, (ii) faixa etária, (iii) escolaridade, (iv) Classe gramatical do elemento não nuclear, (v) Formação do plural no elemento nuclear, (vi) Tonicidade do elemento não nuclear, (vii) Número de sílabas do elemento não nuclear e (viii) Número de sílabas do elemento nuclear.*

3.1. Grupos de fatores sociais

Sexo/gênero

Homens e mulheres diferenciam-se não somente por características fisiológicas (anatomia, timbre e altura de voz – por exemplo), mas também por seus usos linguísticos. Nesse sentido, não se trata de uma diferenciação biológica, mas socialmente construída, uma vez que a consistência entre uso do padrão e conservadorismo por parte das mulheres emerge principalmente em estudos de comunidades de fala ocidentais, que compartilham valores estabelecidos pela natureza das relações socioculturais.

Segundo Labov (1982), as mulheres são mais sensíveis a noções de prestígio de formas linguísticas. De certa forma, atribui-se isso ao maior “formalismo” associado aos papéis sociais femininos. Acerca de processos de mudança, considera Paiva (1992:71):

Quando se trata de implementar na língua uma forma considerada prestigiada, as mulheres tendem a liderar o processo de mudança. Quando, ao contrário, se trata de implementação de uma forma desprestigiada, as mulheres assumem uma atitude conservadora e os homens tomam a ponta do processo de mudança.

Sexo/gênero tem se mostrado relevante nos mais diversos estudos de variação e mudança. Labov (2008[1972]), ao estudar o *r*-pós vocálico no inglês de Nova York, confirma que as mulheres são as que mais utilizam a variante de prestígio. Oushiro (2011), num estudo das *interrogativas-Q* (“o que/ o que que você faz?” *versus* “você faz o que?”), não evidenciou correlação direta entre as variantes e a variável sexo/gênero. Em Rocha (2012), não se verificou correlação entre sexo/gênero e a variação na estrutura da *negação* no português paulistano (com negação pré-verbal, pós-verbal ou pré- e pós-verbal). Tais estudos, conforme comumente se observa na literatura, mostram que a correlação entre sexo/gênero e variáveis morfossintáticas, quando ocorre, é indireta. Entretanto, embora a variável aqui estudada seja morfossintática, é comum observar-se correlação com sexo/gênero (talvez pela importância das diferenciações na base de saliência fônica, conforme se comenta no capítulo 1 e conforme se verifica nos resultados aqui obtidos. Entre os estudos sobre concordância nominal, Scherre (1988, 1994, 1996), com dados do Rio de Janeiro, Salomão (2010) e Fiamengui (2011), com dados de interior paulista, mostram que CN- $\emptyset$ é relativamente menos frequente entre as mulheres do que entre os homens.

No presente estudo, o interesse do grupo sexo/gênero está em verificar se a hipótese de que CN- $\emptyset$, enquanto forma não padrão, é relativamente mais frequente na fala dos homens do que na das mulheres, tanto para alagoanos quanto para paulistanos.

Faixa etária

A disposição dos informantes em três faixas etárias (mais jovens, adultos e mais idosos – como normalmente se faz nas amostras de comunidades de fala enfocadas

nos estudos sociolinguísticos) permite definir uma escala em “tempo aparente”, um construto baseado na hipótese clássica (Bailey 2003) de que a aquisição da linguagem (em termo amplos, sociolinguisticamente falando) se completa por volta dos 15 anos. De acordo com Labov (2008 [1972]), um processo de mudança em curso pode ser vislumbrado (ainda que com dados sincrônicos), quando uma forma variante é menos frequente entre os mais idosos, e consecutivamente mais frequente, em termos relativos, nas outras duas faixas etárias. Assim, no caso de CN-Ø, seria possível falar em mudança em progresso caso tal forma se revelasse mais frequente na fala dos mais jovens, relativamente aos mais velhos.

A fim de verificar, nesta pesquisa, se a idade tem algum peso nas “escolhas” linguísticas dos falantes das subamostras em foco e se o que está ocorrendo é uma mudança em progresso ou variação estável, as faixas etárias estabelecidas são: 1 (20 – 34 anos), 2 (35 – 59 anos) e 3 (60 ou mais anos).

Escolaridade

A relevância do grupo escolaridade nos trabalhos sociolinguísticos reside no fato de que a escola pode exercer influência na dinâmica social de uma variável linguística, sobretudo quando ela é sensível a quanto tempo um falante permaneceu nos bancos escolares – com é o caso da concordância verbal (Scherre e Naro, 1998) e nominal. No geral, verifica-se que, quanto mais são os anos de escolaridade, mais um grupo de falantes tende a empregar variantes prescritas pela norma. Especificamente acerca dos casos de concordância, Scherre (1989:236) afirma:

“inquestionavelmente, as pessoas mais escolarizadas, mais sensíveis às marcas de prestígio e que exercem profissões de trato público tendem a fazer mais concordância e, se não a fazem, são criticadas por nós, que também deixamos de fazer concordâncias verbais e nominais, de forma regular, quer queiramos ou não, quer reconheçamos ou não.”

Nessa pesquisa, o projeto inicial era de se analisar dois níveis de escolaridade “ensino médio” e “ensino superior”, uma vez que a amostra paulistana disponibilizada pelo Projeto SP-2010 é assim organizada. No entanto, conforme já se reiterou, não foi possível encontrar alagoanos em São Paulo que pertencessem à terceira faixa etária e que tivessem com ensino superior. De acordo com a descrição feita no capítulo anterior, foi necessária uma reformulação dos fatores desse grupo, que passaram a ser Ensino fundamental (até 8ª série) e Ensino Médio.

3.2. Grupos de fatores linguísticos

Classe gramatical do elemento não nuclear

Scherre (1988, 1996) analisou o grupo de fatores classes gramaticais não-nucleares associado à posição que ocupavam dentro do SN, uma vez que seus dados incluíram tanto SNs simples (dois elementos constituintes) quanto SNs complexos (três ou mais elementos constituintes). Em seus trabalhos, foram analisadas as seguintes classes (tratando-se do elemento não nuclear): *adjetivo*, *quantificador*, *possessivo*, *artigo*, *demonstrativo* e *indefinido*. Em sua análise chamada de “atomística”, Scherre mostra que qualquer que seja a classe gramatical, o elemento não nuclear apresenta o maior índice de marcas de plural, relativamente ao elemento imediatamente seguinte e ao elemento ao núcleo do SN. Mostra ainda que as classes que mais favorecem a ausência de plural são os adjetivos e substantivos.

No presente trabalho, como as análises se restringiram a SNs constituídos de dois elementos, o interesse está em verificar se há alguma classe morfológica que favorece CN-Ø e se essa tendência se verifica tanto entre alagoanos quanto entre paulistanos. Considera-se que, no caso de SNs plurais simples, o elemento não nuclear é categoricamente plural (de fato, não há casos, nas amostras aqui analisadas, de SNs do tipo “a casas”, em que há marca de plural no elemento nuclear, mas não no elemento à esquerda).

Em primeira instância, os elementos não nucleares (à esquerda) foram codificados de acordo com todas as classes verificadas nas ocorrências extraídas das amostras,

conforme lista o Quadro 5, abaixo. Após a verificação das distribuições dos dados (ver capítulo 4), amalgamações foram feitas, a fim de melhorar o modelo analítico, em termos estatísticos.

Fatores codificados	Exemplos de Ocorrências
o(s), a(s) O	as câmaraØ de caixa eletrônico ontem (es)tava...
Um (ns) uma (s) U	umas quadras pra baixo... mas ela era (menor)... digamos ...
Numerais N	onze casas... onze a doze
Esse(s) essa(s) E	essas coisas... tem muito... é muito cinema se (vo)cê ...
Meu(s) Minha(s) M	minhas amigaØ não sei o quê... mas por outro...
Mesmo(s) Mesma(s) O	mesmas palavras ...meu você está ...
Bastante, melhor (es) B	bastante amigos aí que a gente toma uma cervejinha ...
Certo(s) certa(s) c	certas coisas então que(r) dize(r) ((vo)cê fica ali) apesar ...
Adjetivos d	grandes negócioØ hoje... como é a igreja católica mas ...
Pronomes possessivos K	seus amigos) você... até no trabalho acho que isso ...
Muito(s) muita(s) m	muitos comércios não tem é [hes.] muita coisa porque ...
Algum(ns) Algun(as) a	Alguns nordestinos irrita um pouco sim...
Pouco(s) pouca(s) p	poucas casaØ... mais terreno do que casa...
Quanto(s) quanta(s) Q	quantos anoØ da Philips
Todo(s) toda(s) T	todas juntas... brincavam...
Tanto(s) tanta(s)	Tantas opções de i(r) em show... de i(r) numa ...
Vários, várias V	Várias opçãoØ de lazeR lá
Preposição + artigos W	nos cursoØ vão... (vo)cê vai se apeRfeiçãoando né?
Preposição + demonstrativos X	Nessas alturaØ nós tínhamos uma criação de cabra porco
Diversos, diversas D	diversas atividades tenta(r) rompe(r) com ...
Substantivos S	Pessoas sérias que acham que/ eu acho que elas ...
Aquele(s) aquela(s)	aquelas guerrinhaØ de mamona também
Outro(s) outra(s)	Outros bairros por exemplo Moema (vo)cê faz tudo ...

Quadro 5: Codificação inicial dos fatores do grupo Classe gramatical do elemento não nuclear

Ulteriormente, esses fatores foram recodificados de acordo com as classes e exemplos a seguir:

NUMERAIS

- (24) “o meu minha mãe devia te(r) ah minha mãe acho que tinha mais de TRINTA ANOS quando caso(u) meu pai também e da família da minha mãe foi a única que caso” (SPM3M)
- (25) “...casa de fulano...se ele faz CINCO VIAGEMØ num dia ele já ganha QUINHENTOS REAIS” (ALF3S)

PRONOMES POSSESSIVOS

- (26) “em outro bairro...mas eu acho que não viu? porque aqui... fui criado aqui MEUS PARENTEØ... MEUS AVÔØ... MEUS TIOØ minha família toda são lapiano entende? e acho que dificilmente eu sairia da Lapa”(SPM3F)
- (27) “Ah...também como toda escola pública tinha SUAS PENDÊNCIAS...mas foi tranquilo não tenho muito do...” (ALF1S)

QUANTIFICADORES INDEFINIDOS

- (28) “ passa(r) em pontos estratégicos ...em ALGUMAS RUAS do bairro ...você percebe que a maioria..”. (SPF2S)
- (29) “eu nasci Mata Grande...sabe...é POUCOS HABITANTES...pequeninha mas lá na Paraíba dessa vez...” (AL2FM)

ARTIGOS

- (30) “olha por incrível que pareça o professo(r) que mais me marcou foi no ano passado que deu UNS PROBLEMAS aí eu não consegui termina(r) a minha escola a tempo...” (SPF1M)
- (31) “mais coladinha...e AS MENINAØ do centro gostam de mais de camiseta...assim...eu tirando UNS PARENTEØ...”

PRONOMES DEMONSTRATIVOS

(32) “eu e minha mãe ficamos doente meu pai trabalhava na época né... aí eu tinha medo... porque já tinha né ESSAS PESSOAØ aí no (xxx)... e eu queria volta(r) pra lá mas aí... nós fomos acostumando...” (SPF2M)

(33) “negócio de eh...de samba...então ESSAS ROUPAS sempre pede...esse estilo de música pede roupas...” (ALF2M)

SUBSTANTIVOS

(34) “{é eles pagavam muito} bem ainda pagavam estadia... em HOTÉIS BONS... foi ótimo viu?” (SPF3S)

(35) “desalojados continuam em ESCOLAS MUNICIPALØ e com o início de ano letivo temem não ter para...” (ALF2M)

Nesses casos, entende-se que o núcleo ocupa a posição esquerda do SN. De todo modo, tal posição continua sendo aquele que é categoricamente marcada com plural – ou seja, o elemento que ocupa a posição à direita (normalmente, um adjetivo), é que vai variavelmente concordar em número com o substantivo à esquerda.

ADJETIVOS

(36) “aí eu implorei pra eles pra me mudarem de escola no ensino médio aí eles arranjaram uma escola pra mim no bairro do Ipiranga que aí foi aí eu tenho ÓTIMAS RECORDAÇÃOØ era fantástica até o nível assim de ensino era muito melhor...” (SPF2S)

(37) “não tinha sonho assim de GRANDES COISAØ...nunca tinha visto nem uma televisão quando sai de lá...” (ALF3M)

A fim de garantir a comparabilidade entre as análises dos dados alagoanos e paulistanos, novas amalgamações foram feitas, conforme se mostra no capítulo 4.

Formação do plural no elemento à direita

Este grupo de fatores foi testado e sua relevância confirmada para a ocorrência de CN-Ø nos estudos de Scherre (1988), Campos e Rodrigues (2002), Fiamengui (2011)

e Oushiro (inédito). O que está na base desse grupo é o conceito de saliência fônica (Naro & Scherre 1992), pelo qual se considera que, em alguns casos, a diferença entre a forma singular e a forma plural é mais saliente (porão/porões, azul/azuis) do que em outros (casa/casas). A hipótese, portanto, é que a não realização da concordância de número tenderia a ocorrer mais nos casos menos salientes – ou seja, aqueles em que a formação do plural é regular. Mais uma vez, o interesse aqui é verificar se essa hipótese se confirma em ambos os subgrupos de habitantes da cidade de São Paulo que este trabalho compara.

PLURAL regular

(38) “é mas é/... o bom é o seguinte... se pega mais bandido assim briga ESSAS COISAØ que aparece na televisão... deve se(r) com uma caneta com alguma coisa” (SPF2M)

(39) “é a questão ... de...das mães que cuidam sozinhas DOS FILHOS muitas vezes...não tem com quem deixa(r) eles...” (ALF2S)

PLURAL EM “IS”

(40) “...anda(r) a pé de você anda(r) no ônibus de você... para(r) NOS FARÓIS hoje em dia certo? eu não sei antigamente eu sentia mais segurança... nessa...” (SPF3F)

(41) “nos tivemos que ergue(r) OS MÓVEIS para limpa(r) tudo...a geladeira...o forno...minha irmã até veio me ajudo(u)” (ALF1M)

PLURAL EM ÕES

(42) “Olha...nós sempre assistíamos...assim...Cinderela...Branca de Neve os SETE ANÕES...Chapeuzinho vermelho...ah...era muito gostosa aquela época...” (SPM3M)

(43) “né ... por que eu tinha MUITAS OBRIGAÇÕES.. ai eu tinha que da(r) conta de todas... se eu tivesse uma festa...” (ALF3M)

PLURAL EM “ES”

(44) “...ligado a isso então (gostava) desse tipo de coisa a gente sempre (es)tava em outros bairro (xxx) OUTROS LUGARØ participava aí com o pessoal” (SPM2M)

(45) “aqui não a gente sempre tinha OS PROFESSORES eu julgo que eram bons porque cobrava...” (ALM2S)

Tonicidade do elemento nuclear

Tal grupo foi introduzido nos estudos sobre concordância por Guy (1981), que alega ser raro o efeito da supressão de marca de plural em sílabas tônicas (recordação/recordações). Os trabalhos de Scherre (1988) e Fiamengui (2011) confirmaram a relevância desse grupo de fatores. Na presente pesquisa, entende-se que os oxítonos ou monossílabos (ré, pais, país, feijão e principal), por terem acento na sílaba que vai receber o morfema de plural, favoreça mais a regra de concordância que os paroxítonos e proparoxítonos. Este grupo se constitui dos seguintes fatores:

MONOSSÍLABOS TÔNICOS

(46) “gosto...gosto é...preferiria...eu gosto muito de trabalha(r) com AS MÃOS... eu gosto muito de cria(r) de mexe(r)” (SPM2S)

(47) “vem para a escola para brinca(r)... OS PAIS não falam uma palavra em português...você não sabe o sacrifício...” (ALM2S)

OXÍTONOS

(48) “TANTAS CONSTRUÇÕES... e eu morava naquela baixada onde passa” (SPM3M)

(49) “toda vida usei hoje tenho ALGUMAS MODIFICAÇÕES porque as coisas mudaram...” (ALF3M)

PAROXÍTONOS

(50) “sai fora né tem que se(r) tudo refratário isolante que eles chamam isola OS ISOLANTES para isola(r) o calor para o calor não sai(r) então é feito” (SPF3S)

(51) “...expor ao ridículo de...de usa(r) OS TERMOS do paulista se você não...acho que a sua cultura...” (ALM3S)

PROPAROXÍTONOS

(52) “bom... não ele não fazia... ele trabalhava na (xxx) ele trabalhava em DUAS GRÁFICAØ...meu pai ele sempre foi ele sempre trabalhou com gráfica...” (SPF2M)

(53) “candeeiro...era acende(r) vela se ia pra escola a gente andava CINCO QUILOMETROS a pé...” (ALM2M)

Número de sílabas do elemento não-nuclear

Este grupo de fatores objetiva testar se, nos SNs de dois elementos, itens não-nucleares com maior número de sílabas – ou seja, que contêm mais material fônico – acabam favorecendo a não marcação do plural no elemento nuclear. A hipótese, de acordo com Scherre (1998), é que nas palavras de três, quatro ou mais sílabas é menos frequente ainda deixar de realizar a concordância.

UMA SÍLABA

(54) “nossa eu fiz UNS RESTRATOS em frente aquela igreja super antiga que tem lá em Parati eu fiquei na pousada do...do neto do Dom Pedro” (SPM3M)

(55) “eu creio que...AS PESSOAS nunca deve para(r) de trabalha(r)...só se a pessoa não tem...” (ALM2M)

DUAS SÍLABAS

(56) “então quando eu tinha quinze eles tinham oitenta... então você imagina dois... eu tenho IRMÃOØ GÊMEOS você imagina dois causando num prédio” (SPF1M)

(57) “Suzano... morei OUTROS MESES na Cachoeirinha e depois vim mora(r) no Campo Limpo...” (ALF2M)

TRÊS SÍLABAS

(58) “A gente fazia AQUELAS GUERRINHAØ de mamona também ai que gostoso... como era boa aquela época a gente se divertia muito... muito mesmo...” (SPM2F)

(59) “tem muita afinidade com esses canto(r)... CANTORAS POPULA(R) de hoje da MPB então...” (ALF2M)

QUATRO OU MAIS SÍLABAS

(60) “a segurança eu acho que é problemática...tem POLICIAIS ENVOLVIDOS a gente sabe...não dá pra confiar...” (SPF3F)

(61) “esclarecido quanto...eh...a sociedade ACONTECIMENTOS SOCIAIS então é.. tem mais conteúdo tem mais...” (ALF1S)

Esse grupo de fatores provavelmente não é perfeitamente ortogonal em relação ao grupo “classe gramatical do elemento à esquerda”. Por exemplo, não há artigos que não sejam monossilábicos. Nesse caso, conforme se recomenda para as análise de dados com o Goldvarb X (Cf. Sankoff; Tagliamonte; Smith, 2005), não se podem incluir ambos os grupos num mesmo conjunto de análises estatísticas. Isso acontece também com outros pares de grupos de fatores.

Número de sílabas do elemento nuclear

A hipótese aqui é semelhante à anterior. De acordo com trabalhos anteriores, (sobretudo Scherre, 1989, Fiamengui, 2011), quanto maior o número de sílabas de um elemento do SN, maior a probabilidade de manutenção das marcas de marcas plural.

UMA SÍLABA

(62) “Eu tinha UNS TIOØ que morava aqui...(en)tendeu? Mas depois de um tempo nós perdemos o contato” (SPM2F)

(63) “nasceu no Darcy Vargas...na mão...DAS MÃOS e não teve jeito e morreu ela nasceu com um problema...” (ALM3F)

DUAS SÍLABAS

(64) “Olha...eu sempre visito MEUS IRMÃOØ que mora em Suzano” (SPF2F)

(65) “Assim eu não não...eu creio que tenha sim OUTROS TIPOØ de laze(r) mas eu nunca...nunca fui não...” (ALF1M)

TRÊS SÍLABAS

(66) “subia na rua João... na...perto DOS CÔRREGOØ tinha um campo de futebol...” (SPM3F)

(67) “negócio de Estados Unidos... Paris... ESSES NEGÓCIOS... mas aqui no Brasil acho que o nordeste...” (ALF1M)

QUATRO OU MAIS SÍLABAS

(68) “pra entra(r) pra mora(r)... então OS BURAQUINHOS que ficavam era alegria né?” (SPF2M)

(69) “homem...nem nem sei dize(r) a você porque PROS ESTRANGEIROØ qualque(r) luga(r) que eu podia i(r)...que...” (ALM3F)

Capítulo 4. Resultados

Neste capítulo, são apresentadas as distribuições gerais dos dados obtidos a partir das entrevistas com alagoanos e paulistanos e os resultados de três análises quantitativas realizadas com as ocorrências de SNs de dois elementos extraídos das subamostras. Após explicitar a distribuição geral de dados, reportam-se os resultados de três conjuntos de análises, conforme o quadro de grupos de fatores apresentado ao final do capítulo 3.

No primeiro conjunto de análises, foram incluídos todos os grupos de fatores sociais e linguísticos: *(i) sexo/gênero, (ii) faixa etária, (iii) escolaridade, (iv) classe de palavra do elemento não nuclear, (v) formação de plural do elemento nuclear, (vi) tonicidade do elemento nuclear, (vii) número de sílaba do elemento não nuclear e (viii) número de sílaba do elemento nuclear.*

O segundo conjunto de análises foi necessário devido ao fato de que os grupos *classe de palavras do elemento nuclear* e *número de sílabas do elemento não nuclear* estarem em interação (ou seja, por não funcionarem de maneira independente entre si). Os grupos de fatores incluídos nessas análises foram: *(i) sexo/gênero, (ii) faixa etária, (iii) escolaridade, (iv) número de sílabas do elemento não nuclear e (v) formação de plural do elemento nuclear.*

No terceiro e último conjunto de análises foram incluídos os fatores: *(i) sexo/gênero, (ii) faixa etária, (iii) escolaridade, (iv) número de sílabas do elemento não-nuclear e (v) número de sílabas do elemento nuclear.* Neste caso, a intenção é verificar se a hipótese de que quanto “menores” as palavras que ocupa as duas posições do SN (em termos de quantidade de material fônico), maior a probabilidade de ocorrência de CN-Ø.

4.1. Distribuição geral de dados

Inicialmente, verificaram-se as distribuições gerais de dados, considerando-se sintagmas nominais simples (aqueles com dois elementos), sintagmas nominais

complexos (aqueles com mais de dois elementos) e estruturas predicativas, tanto para alagoanos quanto para paulistanos. Tais distribuições encontram-se nas Tabelas 3 e 4 e figura 3.

Tabela 3: Distribuição geral dos dados alagoanos

	ALAGOANOS				
	CN-∅		CN realizada		Total
	Ocorrência	%	Ocorrência	%	
SNs Simples	688	27	1903	73	2591
SNs Complexos ¹⁷	65	11	445	89	510
Predicativos	59	14	356	86	415

Tabela 4 Distribuição geral dos dados paulistanos

	PAULISTANOS				
	CN-∅		CN realizada		Total
	Ocorrência	%	Ocorrência	%	
SNs Simples	562	22	2045	78	2607
SNs Complexos	28	7	359	93	687
Predicativos	56	12	391	88	447

Figura 3: Percentuais de CN-∅ para paulistanos e alagoanos em SNs simples e complexos, e predicativos.

¹⁷ Foram considerados nesta pesquisa como sintagmas nominais complexos sem concordância aqueles cuja ausência de marcas plural não é realizada em pelo menos um dos elementos do sintagma.

São 2591 as ocorrências de sintagmas nominais plurais “simples” (constituídos de dois elementos) extraídas das 24 entrevistas com alagoanos, das quais 27% são representadas pela não realização da concordância de número (CN-Ø). Das 24 entrevistas com paulistanos, obtiveram-se 2607 ocorrências de sintagmas desse tipo, com 22% de CN-Ø.

Nos sintagmas mais complexos (mais de dois elementos) e nas construções predicativas, as tabelas acima mostram que são aproximadamente cinco vezes menores os totais de ocorrências (respectivamente, 510 e 415), bem como a taxa de não realização da concordância (respectivamente, 11% e 14%). Como o objetivo central aqui é fazer uma comparação entre alagoanos em São Paulo e paulistanos nativos – do ponto de vista da variação na concordância nominal de número – este trabalho se delimita aos casos de sintagmas nominais simples. Oushiro (inédito) oferece uma análise mais abrangente desse caso de variação na fala de paulistanos, tanto da perspectiva do tamanho da amostra quando da profundidade (além de abarcar também SNs com mais de dois elementos, desenvolve também análises atomísticas, a exemplo de Scherre 1988 e 1997).

Considerando-se apenas os SNs de dois elementos, a distribuição dos dados acima permite dizer que a não realização da concordância (CN-Ø) tem frequências semelhantes nos dois subgrupos, mas é um pouco mais baixa entre paulistanos nativos (22%) do que entre os alagoanos (27%). Interessa a este trabalho verificar se esta diferença é significativa e se, em cada grupo, as correlações com grupos de fatores sociais e linguísticos se assemelham ou se diferenciam.

4.2. Análise 1

Os dados de SNs simples foram analisados quantitativamente com a versão mais recente do GoldVarb X (Cf. Sankoff; Tagliamonte; Smith, 2005). Além da variável dependente (com foco na CN-Ø), incluíram-se nove variáveis independentes (ou grupos de fatores - ver capítulo 3). Foram selecionados como estatisticamente

relevantes: (i) *sexo/gênero*, (ii) *faixa etária*, (iii) *escolaridade*, (iv) *classe do elemento não-nuclear* e (v) *formação do plural*. As distribuições de ocorrências e pesos relativos estão organizados na tabela 3. As subamostras de alagoanos e paulistanos foram analisadas separadamente.

Tabela 5: Análise 1 - CN- $\emptyset$ entre alagoanos

ALAGOANOS					
GRUPO DE FATORES	FATORES	TOTAL	%	P.R.	RANGE
FAIXA ETÁRIA	3	304/540	36	.633	33
	2	123/886	12.2	.30	
	1	261/477	35.4	.632	
FORMAÇÃO DO PLURAL	Em “S”	630/1582	28.5	.52	16
	Outros	58/321	15.3	.36	
ESCOLARIDADE	Fundamental	231/394	37	.61	15
	Médio	457/1509	23.2	.46	
SEXO	Masculino	434/1074	28.8	.53	7
	Feminino	254/829	23.5	.45	
CLASSE (ELEM. ESQUERDO	Demonstrativos	493/1267	28	.52	7
	Outros	193/636	23.4	.45	
TOTAL		688/1903	27		
Input: 0.238			Significance: 0.009		

Tabela 6: Análise 1 - CN- $\emptyset$ entre paulistanos

PAULISTANOS					
GRUPO DE FATORES	FATORES	TOTAL	%	P.R.	RANGE
ESCOLARIDADE	Fund.	243/329	42	.74	32
	Médio	319/1746	16	.42	
FORMAÇÃO DO PURAL	Em “S”	527/1711	23.5	.54	24
	Outros	35/334	9.5	.30	
FAIXA ETÁRIA	3	308/963	38	.57	18
	2	120/689	15.3	.39	
	1	134/663	17	.52	
CLASSE (ELEM. ESQUERDO)	Demonstrativos	387/1289	23.1	.53	8
	Outros	175/759	18.7	.45	
SEXO	Masculino	334/841	28.8	.53	7
	Feminino	128/1204	23.5	.46	
TOTAL		562/2045	22		
Input: 0.181			Significance: 0.003		

Nesse primeiro conjunto de análises, todos os grupos de fatores sociais foram selecionados como estatisticamente relevantes, tanto para alagoanos quanto para paulistanos, de modo que a variação na realização da concordância nominal está bastante atrelada aos fatores sociais, já que três dos grupos de fatores linguísticos (tonicidade do elemento nuclear, número de sílabas do elemento nuclear e número de sílabas do elemento não nuclear) não foram selecionados. O total de dados foi submetido a outras duas análises, chamadas de análise 2 e análise 3 que inclui diversas etapas que serão apresentadas nas seções 4.4 e 4.5, nas quais, se atenta para a atuação dos grupos de fatores não selecionados nesta primeira análise.

As tabelas acima mostram que os mesmos grupos de fatores foram selecionados tanto na análise da amostra alagoana quanto na da paulistana. Além disso, para cada grupo selecionado, os fatores com pesos mais altos e mais baixos também coincidem nas duas análises. Entretanto, a ordenação dos grupos nas tabelas mostra que os ranges não coincidem. Por exemplo, entre os alagoanos, o grupo com maior range é faixa etária, enquanto que para paulistanos é a escolaridade. Em ambos os grupos, a formação do plural é o segundo grupo com maior range, mas o terceiro é escolaridade para alagoanos e faixa etária para paulistanos. As interpretações para os resultados são desenvolvidas abaixo, começando-se pelos fatores sociais.

Fatores sociais

De acordo com as tabelas 5 e 6, os grupos de fatores sociais revelaram-se, todos, estatisticamente relevantes, tanto na amostra alagoana quanto na paulistana, evidenciando a importância de sua correlação com a variável em foco. Os resultados detalham-se a seguir.

Faixa etária

Conforme vimos anteriormente, este foi o grupo de fatores com maior range entre os alagoanos e o segundo entre os paulistanos – o que indica sua forte influência na realização da concordância de número no SN, em ambas as amostras.

Tabela 7: CN-Ø e Faixa Etária

		ALAGOANOS				PAULISTANOS			
GRUPOS	FATORES	TOTAL	%	P.R.	R.	TOTAL	%	P.R.	R.
FAIXA ETÁRIA	3	304/540	36	.633	33	308/963	38	.57	18
	2	123/886	12.2	.30		120/689	15.3	.39	
	1	261/477	35.4	.632		134/663	17	.52

Figura 4: CN-Ø e Faixa Etária para alagoanos e paulistanos (P.R.)

Tanto entre os alagoanos quanto entre os paulistanos, os informantes da terceira e da primeira faixa etária se assemelham no favorecimento de CN-Ø, enquanto os de segunda faixa etária desfavorecem tal variante. Para ambos, portanto, não há indicativo, em tempo aparente, de mudança em curso.

A interpretação para esse tipo de resultados normalmente vai na direção de considerar que os informantes da primeira e da terceira faixa etária tendem a ser menos pressionados pelo mercado de trabalho do que aqueles de segunda faixa etária, que se veem mais fortemente cobrados no sentido de empregar formas linguísticas consideradas “corretas” ou aquelas que são socialmente mais bem avaliadas (Scherre, 1989, 1991, 1996, Scherre e Naro, 1998). É interessante observar que, para os informantes mais idosos e mais jovens, é maior entre alagoanos do que entre paulistanos a tendência de favorecimento de CN-Ø (os pesos relativos para alagoanos

são mais altos, e os pesos para paulistanos ficam mais próximos do ponto neutro). Por outro lado, para os informantes de segunda faixa etária, o desfavorecimento de CN-Ø é maior entre alagoanos. De modo geral, portanto, a hipótese do mercado linguístico é mais fortemente evidenciada entre alagoanos.

Em Oushiro, (inédito) que trabalha com a mesma variável numa amostra maior (118 entrevistas sociolinguísticas com paulistanos), constatou-se que este grupo social não foi selecionado como estatisticamente relevante para a concordância nominal, tanto para uma análise atomística quanto para SNs.

Tabela 8: CN-Ø e Faixa Etária para análise atomística e SNs em Oushiro (inédito)

Faixa etária	Atomística			R.	SNs			R.
	P.R.	% CN-Ø	N Total		P.R.	% CN-Ø	N Total	
20-34 anos	[.51]	13,5	6.233	14	[.48]	14,3	5.613	14
35-59 anos	[.42]	11,0	7.173		[.44]	12,2	6.405	
60+ anos	[.56]	17,9	6.478		[.58]	19,2	5.848	

A autora verifica que a concordância nominal parece ser um caso de variação estável em São Paulo. Comparando os resultados obtidos em seu trabalho com os obtidos para paulistanos na presente análise, de fato observa-se praticamente um mesmo padrão; ou seja, embora este grupo não tenha sido selecionado em Oushiro (inédito), os pesos relativos estão muito semelhantes aos obtidos aqui.

Figura 5: CN-Ø e Faixa Etária em duas pesquisas com amostras paulistanas

Escolaridade

A literatura sobre a variação na concordância nominal de número no Brasil vem confirmando a hipótese de que, quanto mais escolarizado o informante, menor é a tendência de não realizá-la (ou seja, menos frequente é CN-Ø). Essa hipótese se confirma também no presente trabalho, tanto entre alagoanos quanto entre paulistanos.

Tabela 9: CN-Ø e escolaridade

		ALAGOANOS				PAULISTANOS			
GRUPO	FATORES	TOTAL	%	P.R.	R.	TOTAL	%	P.R.	R.
ESCOLARIDADE	Fund.	231/394	37	.61	15	243/329	42	74	32
	Médio	457/1509	23.2	.46		319/1746	16	42

Figura 6: CN-Ø e escolaridade para alagoanos e paulistanos

Além de lembrar que escolaridade tem um range maior para paulistanos do que para alagoanos, é interessante verificar que, para os informantes com nível fundamental, CN-Ø tende a ocorrer mais entre paulistanos do que entre alagoanos. Isso pode ser justificado pelo fato de não termos o mesmo número de informantes

com ensino fundamental entre os alagoanos (ver detalhamento no cap. 3), já que não temos um informante do sexo feminino com ensino fundamental na primeira faixa etária, sendo essa célula ocupada por uma informante com ensino médio. Já no nível médio, é pequena a diferença entre alagoanos e paulistanos, mas os primeiros deixam de realizar a concordância um pouco mais.

Os resultados de Oushiro (inédito) se equiparam a entes, no que diz respeito aos paulistanos. Devido ao fato de trabalhar com uma amostra muito mais representativa, e por não ser difícil encontrar informantes das três faixas etárias que não tenham concluído o ensino superior, foi possível estratificar informantes em três níveis de escolaridade (i - Ensino Fundamental, ii – Ensino Médio e iii – Ensino Médio). Essa categorização permitiu entrever que a sistematização é gradual, sendo que os informantes que estudaram até o ensino fundamental favorecem fortemente a CN-Ø (P.R. .81); os que estudaram até o Ensino Médio também favorecem (P.R. .64) enquanto aqueles que concluíram o Ensino superior desfavorecem bastante essa variantes (P.R. .35).

Embora nessa pesquisa não tenha sido trabalhado com Classe Social, pode-se questionar se esses resultados tanto para alagoanos quanto para paulistanos se devem à interferência da escolarização dos informantes ou se se trata dos efeitos das posições sociais que ocupam dentro da sociedade, uma vez que quanto maior nível de escolaridade maiores serão as oportunidade de melhores trabalhos e ascensão social. Oushiro (inédito) testou o fator Classe social em seus dados paulistanos e declara que:

“Parece preferível interpretar que as tendências de emprego observadas em diferentes níveis de escolaridade são decorrência do status social dos falantes e das redes sociais que estabelecem com falantes da mesma classe social.”

Tal interpretação pode ser possível também para os resultados aqui reportados, uma vez que os pesos relativos tanto para alagoanos quanto para paulistanos nos dois níveis de escolaridade (Ensino Fundamental e Ensino Médio) são muito próximos daqueles verificados pela autora, sobretudo para Ensino médio, grupo em que alguns informantes possuem Ensino Fundamental completo.

Sexo/Gênero

Conforme revimos anteriormente, são inúmeras as pesquisas sociolinguísticas que verificam que, em geral, pelo menos nas sociedades ocidentais, as mulheres costumam ser mais sensíveis às formas estigmatizadas numa comunidade de fala (Labov, 2008 [1972]). No que toca à concordância de número no SN, Scherre (1994; 1996) mostra que CN-Ø é menos frequente entre as mulheres do que entre os homens. Nesta pesquisa, interessa é verificar se essa diferenciação se verifica tanto entre alagoanos quanto entre paulistanos:

Tabela 10: CN-Ø e sexo/gênero

		ALAGOANOS				PAULISTANOS			
GRUPO	FATORES	TOTAL	%	P.R.	R.	TOTAL	%	P.R.	R.
SEXO/GÊNERO	Masculino	434/1074	28.8	.53	7	334/841	28	.53	7
	Feminino	254/829	23.5	.45		128/1204	16	.46

Figura 7: CN-Ø sexo/gênero entre paulistanos e alagoanos

Também neste trabalho, os resultados se replicam: CN-Ø é menos frequente entre as mulheres (tanto alagoanas quanto paulistanas), relativamente aos homens. Lembre-se ainda que este grupo de fatores tem o mesmo range na análise de ambas as amostras. Considerando-se a comparabilidade entre elas (mesmo tipo de entrevistas, semelhanças sociais entre os informantes das amostras, mesmos grupos e mesmos fatores em cada conjunto de análises), este conjunto de resultados replica duplamente aqueles de trabalhos anteriores.

Oushiro (inédito) também constatou em seu trabalho que os homens são os que mais concorrem para o favorecimento CN-Ø (P.R. .63) enquanto as mulheres para o desfavorecimento (P.R. .46) – resultados esses que se assemelham muito com os da pesquisa, o que pode ser um indício de que variáveis cujas variantes se diferenciam em graus de estima e prestígio, as formas não padrões são comumente favorecidas pelos homens, e que dentro da cidade de São Paulo, tanto entre paulistanos quanto entre alagoanos, este padrão se repete.

4.3. Fatores Linguísticos

Dos cinco grupos de fatores linguísticos, apenas dois foram selecionados como significativos: *formação do plural do elemento nuclear* e *classe gramatical do elemento não nuclear*.

Formação do plural

Inicialmente, esse grupo foi codificado de acordo com cada tipo possível de formação de plural, conforme se descreve no capítulo 3. A distribuição dos dados nestes fatores está apresentada na tabela 11.

Tabela 11: Distribuição dos fatores no grupo Formação do plural

Paulistanos ¹⁸		
FATORES	N. Ocorrências	%
Plural em S <u>Ex. CasaS</u>	2262	86,7
Plural em ES <u>Ex. SaborES</u>	237	9,1
Plural em IS <u>Ex. PinceIS</u>	77	2,9
Plural em EIS <u>Ex. DificEIS</u>	13	0,5

Essa tabela mostra que há muito poucos dados para todos os casos em que o plural não é regular. Dessa forma, esse grupo foi recodificado em apenas dois fatores: *plural regular* (formado com simples acréscimo do morfema plural S), *plural irregular* (que contempla todas as outras formas). Os resultados para este grupo encontram-se na tabela 10 e na figura 8:

Tabela 12: CN-Ø e formação do plural

		ALAGOANOS				PAULISTANOS			
GRUPO	FATORES	TOTAL	%	P.R.	R.	TOTAL	%	P.R.	R.
FORMAÇÃO DO PLURAL	Regular	630/1582	28.5	.52	16	527/1711	23.5	.54	24
	Irregular	58/321	15.3	.36		35/334	9.5	.30	

¹⁸ Nesta tabela, em que só constam os dados paulistanos, visualiza-se uma “má” distribuição de dados que também se observa na amostra alagoana.

Figura 8: pesos para CN-Ø de acordo com a formação do plural

Tanto para alagoanos quanto para paulistanos, a não realização da concordância é favorecida quando o plural é *regular* (.52 e .54, respectivamente) e desfavorecida quando o plural é *irregular* (.36 e .30, respectivamente). Assim, considerando-se que as análises de cada amostra foram feitas separadamente, praticamente não há diferença entre esses subgrupos de falantes na cidade de São Paulo, no que toca aos pesos dos fatores nesse grupo. Além disso, em ambas as amostras, esse é o segundo grupo de fatores com maior range, o que indica que sua relevância (diante dos demais grupos de fatores) também coincide nas duas amostras, já que os mesmos grupos, com as mesmas configurações de fatores, foram incluídos em ambas as análises.

Dessa forma, a hipótese de que a não realização da concordância é mais frequente quando o plural é regular (em outras palavras, que o plural é mais frequentemente marcado no elemento nuclear do SN quanto maior a diferença entre suas formas singular e plural – de acordo com Scherre (1988) e Fiamengui (2011)) é confirmada em mais uma pesquisa, e para dois conjuntos de informantes.

Classe de palavra do elemento não nuclear

Nos estudos revistos no Capítulo 1, a influência desse grupo de fatores não foi analisada isoladamente; Scherre (1988, 1996) e Fiamengui (2011), por exemplo, focalizaram a atuação do grupo de fatores *classe gramatical* em conjunto com o grupo *posição dos elementos dentro do sintagma nominal* (já que lidaram também com sintagmas nominais de mais de dois elementos). Em São Paulo, Oushiro (inédito) também incluiu o grupo *posição*. No presente estudo, porém, tal grupo não foi considerado, pois focalizam-se apenas sintagmas nominais simples (dois elementos), nos quais o primeiro tem sempre o plural marcado enquanto no segundo a marcação é variável.

Inicialmente, este grupo, assim como o anterior, codificou os dados de acordo com cada classe morfológica possível. Algumas classes não se revelaram frequentes nas amostras, de modo que se realizaram amalgamações. Tendo em vista o interesse de comparar as amostras alagoana e paulistana, a reconfiguração que se revelou quantitativamente mais eficiente tem dois fatores para este grupo: *demonstrativos*, que concentram a maior parte dos dados e, *demais classes*. Os resultados para este grupo encontram-se na tabela 13 e na figura 9:

Tabela 13: CN-Ø e classe de palavra do elemento não nuclear

		ALAGOANOS				PAULISTANOS			
GRUPO	FATORES	TOTAL	%	P.R.	R.	TOTAL	%	P.R.	R.
CLASSE DE PALAVRA	Demonstrativos	493/1267	28	.52	7	387/1289	23.1	.53	8
	Demais classes	193/636	23.4	.45		175/759	18.7	.45

Figura 9: CN-Ø e classe de palavra do elemento não nuclear

Conforme mostram as tabelas 5 e 6 no início deste capítulo, este é um dos grupos de fatores com menor range nessa análise comparativa. Contudo, esse grupo permite verificar que há diferenças na tendência de não realizar a concordância a depender da classe gramatical do elemento não nuclear. Isso já foi verificado em vários trabalhos, mas vale lembrar que, em alguns casos, como em Santos (2010) que analisou o português falado em Pedro Miranda (MG), esse grupo de fatores revela o maior range, com peso relativo bastante elevado para os demonstrativos (.74), que portanto favorecem muito CN-Ø. Contudo, na grande maioria dos trabalhos (Poplack 1980; Scherre 1988, 1996, Carvalho 1977, entre outros), os demonstrativos favorecem a realização da concordância.

Nos termos da comparação que este trabalho faz, interessa ressaltar que alagoanos e paulistanos não se diferenciam: além da semelhança entre os ranges desse grupo de fatores nas duas amostras, os pesos relativos também se assemelham.

4.4. Análise 2

No início deste capítulo, menciona-se que alguns dos grupos de fatores interagem entre si (ou seja, não são ortogonais). Num segundo conjunto de análises, diferentemente da primeira, os grupos incluídos foram: (i) *sexo/gênero*, (ii) *faixa*

etária, (iii) escolaridade, (iv) número de sílabas do elemento não nuclear e (v) formação de plural do elemento nuclear. Ou seja, a diferença aqui está no grupo em (iv) – que, aliás, não foi selecionado como estatisticamente relevante. Reitera-se que os dados são diretamente comparáveis, já que os mesmos grupos foram incluídos para cada amostra, com a mesma configuração.

Tabela 14: Análise 2 – CN-Ø para alagoanos

ALAGOANOS					
GRUPOS	FATORES	TOTAL	%	P.R.	R.
FAIXA ETÁRIA	3	304/540	36	.633	33
	2	123/886	12.2	.30	
	1	261/477	35.4	.632	
PLURAL DO ELEM. NUCLEAR	Em “S”	630/1582	28.5	.52	17
	Outros	58/321	15.3	.36	
ESCOLARIDADE	Fund.	231/394	37	.61	15
	Médio	457/1509	23.2	.46	
SEXO/GÊNERO	Masculino	434/1074	28.8	.53	6
	Feminino	254/829	23.5	.45	
TOTAL		688/1903	27		
Input: 0.240 Significance: 0.005					

Tabela 15: Análise 2 – CN-Ø para paulistanos

PAULISTANOS					
GRUPOS	FATORES	TOTAL	%	P.R.	R.
ESCOLARIDADE	Fund.	243/329	42	74	33
	Médio	319/1746	16	42	
PLURAL DO ELEM. NUCLEAR	Em “S”	527/1711	23.5	.54	23
	Outros	35/334	9.5	.30	
SEXO/GÊNERO	Masculino	334/841	28	.53	22
	Feminino	128/1204	16	.46	
FAIXA ETÁRIA	3	380/693	38	.57	20
	2	120/689	15	.39	
	1	134/663	17	.52	
TOTAL		562/2045	22		
Input: 0.183 Significance: 0.000					

Com exceção do número de sílabas do elemento não nuclear, os mesmos grupos de fatores que foram selecionados como estatisticamente relevantes na

Análise 1 também o foram na Análise 2. Os pesos relativos para estes grupos relevantes na presente análise se assemelham aos da anterior, assim como seus ranges.

4.5. Análise 3

Realizou-se também um terceiro conjunto de análises, excluindo-se o grupo Classe de palavra e Formação de plural (relativamente ao conjunto 1) e incluindo-se (i) *sexo/gênero*, (ii) *faixa etária*, (iii) *escolaridade*, (iv) *número de sílabas do elemento não nuclear* e (v) *número de sílabas do elemento nuclear*. O grupo de fatores (iv) foi selecionado como estatisticamente relevante na amostra alagoana, mas não na paulistana.

Tabela 16: Análise 3 – CN-Ø para alagoanos

ALAGOANOS					
GRUPOS	FATORES	TOTAL	%	P.R.	R.
FAIXA ETÁRIA	3	304/540	36	.633	33
	2	123/886	12.2	.30	
	1	261/477	35.4	.632	
NÚMERO DE SÍLABAS DO ELEMENTO NUCLEAR	Monossílabas	61/78	68	.68	22
	Dissílabas	343/883	51	.51	
	Trissílabas	284/942	46	.45	
ESCOLARIDADE	Fund.	231/394	37	.61	15
	Médio	457/1509	23.2	.46	
SEXO/GÊNERO	Masculino	434/1074	28.8	.53	6
	Feminino	254/829	23.5	.45	
	TOTAL	688/1903	27		
Input: 0.261			Significance: 0.009		

Tabela 17: Análise 3 – CN-Ø para paulistanos

PAULISTANOS					
GRUPOS	FATORES	TOTAL	%	P.R.	R.
ESCOLARIDADE	Fund.	243/329	42	74	33
	Médio	319/1746	16	42	
SEXO/GÊNERO	Masculino	334/841	28	.53	22
	Feminino	128/1204	16	.46	
FAIXA ETÁRIA	3	308/693	38	.57	20
	2	120/689	15	.39	
	1	134/663	17	.52	
NÚMERO DE SÍLABAS DO ELEMENTO NUCLEAR	Monossílabas	397/1343	23	[.51]	5
	Dissílabas	56/208	21	[.50]	
	Trissílabas	109/494	18	[.46]	
	TOTAL	688/1903	27		
Input: 0.261			Significance: 0.009		

Os resultados para os grupos sociais nesta análise se assemelham aos das anteriores, tanto para alagoanos quanto para paulistanos. Dos dois grupos de fatores linguísticos, o número de sílabas do elemento não-nuclear não foi selecionada para nenhuma das duas amostras. Já o grupo *número de sílabas do elemento nuclear* foi selecionado apenas para amostra alagoana – CN-Ø é mais frequente quando o elemento nuclear é um *monossílabo* (.68) e gradativa e relativamente menos frequente quando é um dissílabo (.51) ou um trissílabo (.45). Esses resultados confirmam a hipótese testada por Scherre (1988) Fiamengui (2011) de que elementos com menos material fônico favorecem a ausência de marca de plural. Para Oushiro (inédito) este grupo de fatores também não se mostrou significativo para os paulistanos investigados, o que reforça que esta não é mesmo uma variável que se correlacione com a realização do plural no SN entre paulistanos.

Figura 10: CN-Ø e número de sílabas do elemento nuclear (para alagoanos)

4.6. Cruzamentos

Os resultados dos conjuntos de Análises 1, 2 e 3 destacam a relevância dos grupos de fatores sociais para a compreensão dos mecanismos que colaboram para a não realização da concordância nominal. A seguir, o cruzamento entre esses grupos permitem mais algumas considerações.

Sexo/faixa etária

Os percentuais na tabela 18 são organizados em dois gráficos que mostram diferenças nas ocorrências de CN-Ø entre homens e mulheres, das diferentes faixas etárias, nas duas amostras:

Tabela 18: CN-Ø no cruzamento entre Sexo/Gênero e Faixa etária

Faixa etária	Alagoanos				Paulistanos			
	Masculino		Feminino		Masculino		Feminino	
	N	%	N	%	N	%	N	%
1	241/414	52	47/324	15	54/358	15	80/439	18
2	56/530	10	67/366	16	86/390	22	34/419	8
3	164/508	32	140/336	42	194/427	45	114/574	20

Figura 11 – CN-Ø no Cruzamento entre Sexo/Gênero e Faixa etária para Alagoanos

Figura 12 – CN-Ø no Cruzamento entre Sexo/Gênero e Faixa etária para paulistanos

Os resultados indicam que os homens alagoanos mais jovens são aqueles em cuja fala CN-Ø é mais frequente (52% dos casos), enquanto que, entre os paulistanos, esse fato se observa entre os homens mais velhos (45%). Além dessas frequências mais altas de CN-Ø, é interessante observar que, entre os alagoanos, há um padrão que se aproxima daquele que se interpreta como um caso de mudança em tempo aparente para as mulheres, já que CN-Ø é mais infrequente entre as mais jovens (tanto segunda

quanto terceira faixa etária). Já entre os paulistanos, tal padrão se verifica entre os homens, enquanto que entre as mulheres repete-se o padrão de variação estável, descrito antes do cruzamento.

No geral, esses dados permitem dizer que, se nas comparações anteriores, alagoanos e paulistanos mais se assemelham do que se diferenciam, estes cruzamentos permitem visualizar uma distinção importante. Entre os alagoanos, a hipótese do mercado linguístico parece se aplicar aos homens, mas não às mulheres, enquanto que, entre os paulistanos, observa-se o contrário. O fato de que os homens alagoanos mais jovens são aqueles que mais deixam de realizar a concordância permite ainda formular uma questão relativa ao significado social de tal variante em São Paulo. Já que não há trabalhos sobre a variável CN na fala alagoana no estado de Alagoas, e já que ainda não se estudou a relação entre a ocorrência de CN-Ø e o tempo de permanência dos alagoanos em São Paulo, seria importante, num outro trabalho, tentar verificar por que CN-Ø é tão mais frequente entre os alagoanos mais jovens.

Sexo/Escolaridade

Os resultados desses cruzamentos são apresentados na tabela 19 e nas figuras 13 e 14 abaixo:

Tabela 19 – CN-Ø no Cruzamento entre Sexo/Gênero e Escolaridade

Escolaridade	Alagoanos				Paulistanos			
	Sexo/Gênero				Sexo/Gênero			
	Masculino		Feminino		Masculino		Feminino	
	Total	%	Total	%	Total	%	Total	%
E. Fundamental	158/400	40	73/225	32	168/233	72	75/339	22
E. Médio	276/1108	25	181/856	21	166/942	18	153/1093	14

Figura 13: Cruzamento entre Sexo/Gênero e Escolaridade para Alagoanos

Figura 14: Cruzamento entre Sexo/Gênero e Escolaridade para paulistanos

No geral, esses cruzamentos mostram que CN-Ø é sempre menos frequente entre mulheres do que entre homens. Entretanto, essa diferenciação é bastante semelhante, para alagoanos, nos dois níveis de escolaridade. Já para paulistanos, no nível mais baixo de escolaridade, a diferenciação é bem maior do que no nível médio.

Faixa etária e escolaridade

Conforme vimos no Capítulo 2, não há, na subamostra paulistana, nenhum informante de primeira faixa etária com nível fundamental de escolaridade. Contudo, para este cruzamento, operou-se com três níveis de escolaridade – diferentemente do que se fez na apresentação dos resultados até aqui.

Tabela 20 – CN-Ø no Cruzamento entre Faixa etária e Escolaridade

Faixa etária	Alagoanos						Paulistanos					
	Escolaridade											
	Fundamental		Médio		Superior		Fundamental		Médio		Superior	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
3	174/363	48	120/359	33	10/122	8	156/408	38	148/523	28	4/70	6
2	28/194	14	44/215	20	51/600	8	87/164	53	22/298	7	3/347	3
1	29/68	43	191/446	43	41/224	18	0	0	131/578	23	3/129	1

Figura 15: CN-Ø no Cruzamento entre Faixa etária e Escolaridade para Alagoanos

Figura 16: CN-Ø no Cruzamento entre Faixa etária e Escolaridade para Paulistanos

Os resultados dispostos acima mostram que o padrão de distribuição de CN-Ø pelos diferentes níveis de escolaridade, nas diferentes faixas etárias, é bastante mais variável entre os alagoanos do que entre os paulistanos. Entre estes últimos, apesar de não haver nenhum informante com nível fundamental na primeira faixa etária, pode-se dizer que o padrão se repete: nas três faixas, a frequência de CN-Ø cai bastante, quanto maior é o nível de escolaridade. Entre os alagoanos, contudo, esse padrão só é observado, tal e qual, entre os mais idosos; para aqueles da segunda faixa etária, CN-Ø é mais frequente entre os informantes que estão no nível médio do que entre aqueles que do nível fundamental de escolaridade. Por outro lado, entre os mais jovens, não há diferença entre os níveis fundamental e médio, nos quais a frequência de CN-Ø é igualmente alta (48%), em relação à média.

No primeiro conjunto de análises, já havíamos visto que o grupo escolaridade – aquele que teve maior *range* para os paulistanos – foi apenas o terceiro com maior *range* para os alagoanos. Os cruzamentos acima mostra, então, mais uma diferença entre alagoanos e paulistanos, no que diz respeito à realização da CN.

Conclusão

O principal objetivo dessa pesquisa sociolinguística foi verificar como se dá a variação na concordância nominal de número em SNs constituídos de dois elementos, na fala de imigrantes alagoanos e paulistanos na cidade de São Paulo. Os dados foram extraídos de uma amostra balanceada de 48 entrevistas, sendo que aquelas com paulistanos foram selecionadas dentre as coletadas pelo GESOL, enquanto que as alagoanas foram gravadas e transcritas especificamente para esta dissertação.

Das entrevistas com alagoanos, foram extraídas 2591 ocorrências de SNs plurais, das quais 27% são representadas pela não realização de concordância de número (CN-Ø). Das entrevistas paulistanas, obtiveram-se 2607 ocorrências, com 22% de CN-Ø. Os sintagmas complexos (mais de dois elementos) e as construções predicativas não foram analisadas quantitativamente; além de serem em média cinco vezes menos frequentes que os SNs de dois elementos, o objetivo central da comparação entre os dois subgrupos na cidade de São Paulo permitiu que se focalizasse num desses conjuntos de dados – que foram analisados quantitativamente com o pacote estatístico Goldvarb X (Sankoff; Tagliamonte; Smith, 2005).

Todos os fatores sociais incluídos nas análises foram selecionados como estatisticamente significativos. Evidencia-se que, na cidade de São Paulo, pelo menos para os dois grupos de falantes investigados, os fatores sociais são, no geral, mais importantes do que os linguísticos.

A *Faixa etária* foi a variável mais significativa entre os alagoanos (*range* 33) e o segundo entre os paulistanos (*range* 18). Em ambos os grupos, os informantes da terceira e primeira faixa etária se assemelham no favorecimento de CN-Ø, enquanto os de segunda faixa etária desfavorecem tal variante. Não há, portanto, indicativo de mudança em curso em tempo aparente. É interessante observar que, para os informantes mais idosos e mais jovens, é maior entre alagoanos do que entre paulistanos a tendência ao favorecimento de CN-Ø. Por outro lado, para os informantes de segunda faixa etária, o desfavorecimento de CN-Ø é maior entre alagoanos. De modo geral, portanto, a hipótese do mercado linguístico é mais fortemente evidenciada entre alagoanos.

Quanto à hipótese de que quanto menor o nível de escolaridade do informante maior a probabilidade de que favorecimento da CN-Ø, ela é confirmada por mais este trabalho. Para os paulistanos, é nesse grupo que se observa maior diferenciação entre os fatores (*range* 33); já para os alagoanos (entre os quais a *faixa etária* teve o maior *range*), a escolaridade apresenta menor diferenciação (*range* 15). Verifica-se também que, para os informantes com nível fundamental, CN-Ø tende a ocorrer mais entre paulistanos do que entre alagoanos. Já no nível médio, é pequena a diferença entre alagoanos e paulistanos, mas os primeiros deixam de realizar a concordância um pouco mais.

O terceiro grupo de fatores sociais com maior *range* para ambos os grupos de falante foi *sexo/gênero*. Assim como em alguns dos outros trabalhos que investigam a concordância nominal, CN-Ø é menos frequente entre as mulheres (tanto alagoanas quanto paulistanas), relativamente aos homens. Pode-se considerar, dessa forma, que também neste estudo as mulheres se revelam as mais conservadoras quanto ao emprego da variante estigmatizadas.

Os trabalhos realizados por Scherre (1988), Santos (2010), Fiamengui (2011) e Oushiro (inédito) são as referências mais importantes na orientação quanto ao teste de hipóteses linguísticas. A *formação do plural do elemento nuclear*, em razão da pouca quantidade de ocorrências em certos casos, foi recodificada em dois fatores: *plural regular* (formado com acréscimo de -S) e *plural irregular* (todas as demais morfologias). Este grupo foi selecionado como relevante tanto para alagoanos (*range* 16) quanto para paulistanos (*range* 24), sendo que CN-Ø é a desfavorecida quando os plurais são irregulares (P.R. .52 na amostra alagoana e .54 na paulistana). Como os pesos relativos para o favorecimento de CN-Ø em casos de plurais regulares são muito próximos do ponto neutro, tanto para alagoanos quanto para paulistanos, parece mais interessante atentar para o fato de que essa forma é consistentemente evitada, por ambos os grupos, nos casos de plurais irregulares.

Quanto à *Classe gramatical* do elemento à esquerda, conforme discutiu-se anteriormente, trata-se do grupo de fatores que se relaciona à posição das marcas plurais nos elementos do SN – um grupo que não se analisou nessa pesquisa, já que se verificou marcação (nos termos de Scherre 1988) categórica do plural no elemento à esquerda. Assim, a pergunta que justifica o grupo *Classe* é se, a depender da natureza

morfológica do elemento à esquerda (um artigo, um adjetivo, um quantificador indefinido, etc) a concordância seria mais, ou menos, frequentemente realizada no elemento à direita. Em razão das distribuições de dados, foi necessário amalgamar fatores de forma a garantir a comparabilidade das duas subamostras de falantes – *demonstrativos* versus demais classes). Tanto para os alagoanos quanto para os paulistanos, CN-Ø é um pouco favorecida quando o elemento à esquerda é um demonstrativo (pesos .52 e .53, respectivamente). Ressalta-se, assim, que alagoanos e paulistanos não se diferenciam neste aspecto.

Finalmente, o *número de sílabas do elemento nuclear* foi selecionado como relevante apenas para os alagoanos; CN-Ø é favorecida quando o elemento nuclear é um *monossílabo* (.68) e gradativa e relativamente menos frequente quando é um *dissílabo* (.51) ou um *trissílabo* (.45). Esses resultados confirmam a hipótese testada por Scherre (1988) e Fiamengui (2011) de que elementos com menos quantidade de material fônico favorecem a ausência de marca de plural.

Os resultados a que esta pesquisa chega permitem ver que alagoanos e paulistanos residentes na cidade de São Paulo se assemelham bastante: ambos os grupos mais realizam concordância do que deixam de realizá-la (embora os percentuais de CN-Ø sejam um pouco maiores entre os paulistanos). Também se assemelham no que diz respeito ao favorecimento de CN-Ø pelos mesmos fatores nos diferentes grupos. Uma diferença que salta aos olhos diz respeito ao fato de que o grupo com maior range é faixa etária para alagoanos e escolaridade para paulistanos. Além disso, os cruzamentos entre os grupos de fatores sociais mostram distribuições diferentes, principalmente para homens e mulheres alagoanas de diferentes faixas etárias, e para alagoanos mais jovens e mais idosos com diferentes níveis de escolaridade. Se CN-Ø tem algum significado social “especial” para paulistanos (cf. Oushiro, inédito), as análises apresentadas aqui poderiam levar à pergunta sobre se os alagoanos que vivem em São Paulo compartilham, com os paulistanos nativos, percepções – sociolinguisticamente falando – das variantes dessa variável. Essa pergunta, contudo, ultrapassa a delimitação do presente trabalho, que lida com dados de produção. Por outro lado, não se tem notícia de trabalhos sobre CN-Ø desenvolvidos com amostras de alagoanos que vivem nas cidades de que são naturais.

Finalmente, pode-se dizer que, pelo menos no que diz respeito à variação na realização da concordância nominal de número, alagoanos e paulistanos mais se assemelham do que se diferenciam. Estudos futuros, sobre a produção (e percepção) de outras variáveis, poderão se somar a este na discussão de fatores sociais e linguísticos que permitem entender esses dois grupos de falantes como constitutivos ou não de uma mesma comunidade de fala.

Referências bibliográficas

ALVES, Maria Isolete Pacheco Menezes. *Atitudes linguísticas de nordestinos em São Paulo (abordagem prévia)*. Campinas: UNICAMP, 1979.

ANTONINO, V. A concordância nominal em predicativos do sujeito e em estruturas passivas no português popular do interior do estado da Bahia. Dissertação de mestrado. Universidade Federal da Bahia – Campus Universitário Ondina Salvador, 2007, 123f.

BRAGA, M. L. (1977). *A Concordância de número no sintagma nominal no triângulo mineiro*. PUC, Rio de Janeiro, 88p. Dissertação de Mestrado, inédito.

DIAS, Maria Clara Alves Correa. *A variação na concordância nominal: um contraste entre o urbano e o rural na fala brasiliense*. Dissertação de Mestrado. UnB., 1993.

GUY, G. R.. *Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax, and language history*. Philadelphia: University of Pennsylvania, 1981. PhD Dissertation, mimeo.

LABOV, William (1963) *“The social motivation of a sound change.”* Word, vol. 19, 273–309.

LABOV, William (2001). *Principles of Linguistic Change. Social Factors*. Oxford: Blackwell.

LABOV, W. (1972:[2008]) *Padrões Sociolinguísticos*. São Paulo: Parábola. Tradução Brasileira: Marcos Bagno, Maria Marta Pereira Scherre e Caroline Rodrigues Cardoso.

MALINOWSKI, B. Objeto, Método e alcance desta pesquisa. In: Os Argonautas do Pacífico Ocidental. Malinowski, Coleção Os Pensadores. São Paulo: Abril Cultural, 1976.

MENDES, R. B. ; OUSHIRO, L. . O paulistano no mapa sociolinguístico brasileiro. Alfa: Revista de Linguística (UNESP. Online), v. 56, p. 973-1001, 2012.

MILROY, Leslie (2004) Social networks. In: Chambers, J.K., P. Trudgill & N. Schilling-Estes (eds.) The Handbook of Language Variation and Change . Oxford: Blackwell.

OUSHIRO, L; Mendes, R. B. . A pronúncia do (-r) em coda silábica no português paulistano. In: Estudos da Linguagem. Casamento entre temas e perspectivas. 1. ed. João Pessoa: Idéia/Editora Universitária da UFPB, 2011. v. 1. 365 p. da HORA, Dermeval (Org.) ; NEGRÃO, E. V. (Org.)

OUSHIRO, L. ; MENDES, R. B. . Cross-over effects of variable nasal /e/ in Brazilian Portuguese. In: Linguistic Society of America 2013 Annual Meeting, 2013, Boston-MA. Linguistic Society of America Meeting Handbook, 2013. v. 1. p. 190-190

OUSHIRO, L. (2011b) “Identidade na pluralidade: produção e percepção linguística na cidade de São Paulo”. Projeto de doutorado.

OUSHIRO, L. . Tratamento de dados com o R para análises sociolinguísticas. In: FREITAG, Raquel Meister Ko.. (Org.). Metodologia de Coleta e Manipulação de Dados em Sociolinguística. 1ed.São Paulo: Editora Edgard Blücher, 2014, v. , p. 134-177.

ROCHA, Rafael Stoppa. A negação dupla no português paulistano. 2013. Dissertação (Mestrado em Semiótica e Linguística Geral) – Faculdade de filosofia, Letras e Ciências Humanas, Universidade de São Paulo, 2013.

RODRIGUES, Angela C. S. (2009) Fotografia sociolinguística do português do Brasil: o português popular em São Paulo. In: Castilho, Ataliba T. (ed.), *História do Português Paulista*. Campinas: Instituto de Estudos da Linguagem/UNICAMP.

SACCONI, Luiz Antônio. Gramática essencial da língua portuguesa: teoria e prática. 4. ed. São Paulo: Atual, 1989. In: MOLLICA, Maria Cecília & BRAGA, Maria Luiza (orgs). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2003.

SALOMÃO, Mircia Hermenegildo. *A variação de pluralidade nas estruturas predicativas da variedade falada na região de São José do Rio Preto*. Dissertação de mestrado. UNESP, 2010.

SANTOS, Lília Soares Miranda. Sobre a ausência de concordância nominal no português falado em Pedro Leopoldo – MG: Uma abordagem variacionista. Dissertação de mestrado. UFMG, 2010.

SCHERRE, Maria Marta Pereira. *Reanálise da concordância nominal em português*. Tese de doutorado. UFRJ, 1988.

SCHERRE, M.M.P. *A concordância de número nos predicativos e nos participípios passivos*. *Organin*, Porto Alegre. V.d, n.17, 1991

SCHERRE, M. M. P. *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. 12:37-49. dez. de 1994.

SCHERRE, M.M.P. *Sobre a influência de três variáveis relacionadas na concordância nominal em português*. In: OLIVEIRA E SILVA, SCHERRE, M.M.P. (orgs.) In: *Padrões sociolinguísticos*. Análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo brasileiro, 1996a, p.85-118.

SCHERRE, M.M.P.; Naro, A.J. *Sobre as origens do português popular no Brasil*. DELTA, v.9, n, 1, p.01-14, 1993.

SCHERRE, M.M.P.; Naro, A.J. *Concordância de número no português do Brasil: um caso de variação inerente*. In: HORA, D. da (org.) *Diversidade linguística no Brasil*. João Pessoa: Ideia. 1997, p.93-114.

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.

SCHERRE, M. M. P. Paralelismo linguístico. *Revista Estudos Linguísticos* 7(2), p. 29-59, 1998.

SCHERRE, M. M. P. Phrase-level parallelism effect on noun phrase number agreement. *Language Variation and Change* 13, p. 91-107, 2001.

Anexo I: Roteiro das entrevistas

As entrevistas da amostra alagoana foram gravadas seguindo as orientações do Projeto SP2010-Construção de uma amostra da fala paulistana¹⁹. Abaixo seguem as orientações gerais desde o primeiro contato com os informantes até a gravação das entrevistas.

Instruções gerais

Antes da entrevista

Contatar o informante.

- Não entrevistar ninguém próximo do convívio diário (parentes, amigos, colegas de trabalho). Escolher informantes que, na Ficha do Informante, são definidos pelos graus 4 e 5. 1
- Entrevistar “amigos de amigo”: procurar, através de seus contatos, outras pessoas que eles conheçam e que sejam informantes potenciais para esta amostra.
- Entrevistar apenas pessoas que nasceram e foram criadas na cidade de São Paulo.
- Entrevistar pessoas que conhece “de vista”: o “cara” do mercadinho, do açougue, da padaria... posteriormente pode-se pedir a esses mesmos informantes a indicação de outras pessoas para serem entrevistadas. De acordo com experiências prévias, os informantes, após terem participado de uma entrevista, em geral não apresentam resistências a indicar novos informantes.

Conversa preliminar

- Apresentar-se como aluno universitário. Dizer ao informante que está fazendo uma pesquisa sobre vários bairros de São Paulo para saber um pouco sobre como é a vida em cada região da cidade e que o interesse da pesquisa é no dia-a-dia dos paulistanos – ou seja, não é nada teórico, nada elaborado. O objetivo é saber como é a vida no

¹⁹ Processo FAPESP nº 2011/09278-6

bairro e os pontos positivos e negativos da região em que o informante vive e da cidade de São Paulo como um todo.

- Essa conversa prévia também leva à obtenção de informações gerais (p.ex., se, de fato, o informante corresponde a um dos perfis estabelecidos pelo projeto) e ajuda a “quebrar” um pouco o grau de formalidade.
- Dizer que a entrevista dura entre uma hora e uma hora e meia. Perguntar ao possível informante quando é o melhor dia e horário para ele. É melhor não gravar a pessoa no mesmo dia se ela se mostrar pouco disponível. Fazer a entrevista preferencialmente na casa do informante.

Pouco antes da gravação

- Antes de iniciar a entrevista, escolher um local adequado para desenvolvê-la: evitar locais barulhentos e/ou abertos, sugerir que a TV seja desligada (caso esteja ligada), evitar sentar perto de janelas ou outros locais em que pode haver barulho (p.ex., perto de um ventilador). De preferência, escolher um local acarpetado – a sala normalmente é o melhor lugar para se fazer uma entrevista.
- Explicar ao informante os objetivos gerais da pesquisa e obter o seu consentimento para ser gravado.
- Pedir que o informante coloque o microfone em si e, caso seja necessário ajustá-lo, instruir o informante para que ele mesmo o faça ou pedir licença para ajustá-lo. Apesar de o microfone potencialmente aumentar o grau de formalidade da situação, ele é importante para garantir uma boa qualidade de gravação, o que posteriormente facilitará a tarefa de transcrição e possibilitará o desenvolvimento de trabalhos a respeito de variáveis fonético-fonológicas.
- Se sentir que o seu informante está um pouco tenso ou desconfiado, conversar um pouco mais com ele antes de começar a gravação. Responder todas as dúvidas que ele possa ter, deixar claro que sua identidade não será divulgada nos resultados da pesquisa e esclarecer que a conversa pode ser interrompida a qualquer momento. Durante a entrevista
- A regra mais básica é falar POUCO! O objetivo é proporcionar as condições para que o informante fale, de forma mais natural possível, dentro da situação de entrevista sociolinguística. Para tanto:
- Escutar o que o seu informante tem a dizer. Mostrar um interesse genuíno pelo que ele está dizendo. Se ele mencionar alguma coisa interessante/pitoresca, pedir que ele

fale mais: “Nossa! E como foi isso?”... A sua postura, durante toda a entrevista, deve ser a de alguém que está lá para aprender.

- Fazer perguntas curtas. Ao mesmo tempo, evitar fazer perguntas cuja resposta é simplesmente “sim” ou “não”; se for fazê-las, ter outra pergunta na
- ponta da língua para que o informante continue falando – p.ex. “Você gosta de morar aqui? (deixar o informante responder) Por que você escolheu morar nesse bairro?”
- Evitar interromper o informante enquanto ele está falando e evitar muitas sobreposições de vozes.
- Os tópicos abaixo são apenas para a sua organização, mas a conversa deve fluir. Não falar coisas como: “Agora vamos falar sobre a sua família”. As perguntas da primeira parte do roteiro são sugestões para manter o fluxo da conversa e, pela experiência das gravações-piloto, fornecem uma sequência “natural” de tópicos a serem tratados. Ser flexível para explorar temas relacionados que se revelem de interesse para o informante. Por outro lado, as perguntas da segunda parte do roteiro são obrigatórias e devem ser feitas, na medida do possível, na mesma sequência em que são apresentadas.
- As gravações devem durar entre 60 e 70 minutos. Administrar o tempo de forma que dê para cobrir todos os tópicos.
- Se for à casa do informante, posteriormente fazer anotações que julgar interessantes e que possam levar a inferências a respeito da classe social, infraestrutura do bairro (se há uma associação de bairro, locais de lazer) etc. Isso também pode facilitar o contato com novos informantes. Depois da entrevista
- Pedir ao informante que assine a declaração.² Explicar que ela é uma declaração de que ele sabia que estava sendo gravado e que serve para proteger os seus dados pessoais. Deixar claro que nenhum dado pessoal será publicado.
- Pegar informações adicionais que podem não ter surgido durante a entrevista. É melhor coletar os dados da Ficha do Informante depois da gravação – o informante se sentirá mais seguro de que seus dados pessoais não foram gravados.
- Em casa, preencher a Ficha do Informante o mais cedo possível. Vai ser mais fácil encontrar/organizar as informações. Depois provavelmente será mais difícil lembrá-las!

Roteiro da entrevista

I PARTE - BAIRRO (aprox. 10 min.)

Objetivos: descobrir o “grau de enraizamento” do informante no bairro onde vive/outros bairros; descobrir o “grau de mobilidade” da pessoa

1. Há quanto tempo você mora na (Mooca, Pinheiros, Bexiga...)?
2. Você gosta de morar aqui?
3. Por que você escolheu morar neste bairro? (manter em mente que o informante pode não ter “escolhido” morar ali: pode ter sido porque a família já morava lá etc.)
 - a. Se o informante mora há bastante tempo: Como era o bairro antes/quando você se mudou pra cá? Mudou muito?
 - b. Se o informante não mora lá há muito tempo, ou se nem sempre viveu ali: Em que outros lugares você já morou? Como era lá comparado com esse bairro aqui? Onde você preferia morar?
4. Você tem bastante contato com as pessoas aqui do bairro? 5. Hoje, tem algum outro bairro em que você gostaria de morar?

INFANCIA (aprox. 5 min.)

Objetivos: relaxar o informante (em geral, as pessoas gostam de falar de sua infância); obter informações sobre mudanças no bairro/cidade de São Paulo; grau de mobilidade do informante; obter informações sobre escolaridade

6. E como foi a sua infância (no bairro X)? Você pode contar um pouco de como foi, o que você fazia...?
 - a. brincava na rua/dentro de casa? Do que vocês brincavam?
 - b. como eram os seus pais? eram rígidos...?
7. Você foi pra escola no mesmo bairro? Como era a escola? Você sempre estudou na mesma escola? Você gostava de ir pra escola? Tem algum professor que te marcou? Até que série você estudou?
8. Enquanto ainda era criança/adolescente, você ia pra outros lugares dentro da cidade de São Paulo? (pra onde, pra fazer o quê...)

FAMILIA (aprox. 5 min.)

Objetivos: obter informações sobre rede social do informante, grau de enraizamento no bairro/cidade

9. Você tem irmãos? (é possível que esta informação já tenha aparecido na parte sobre infância; neste caso, falar algo como: “Você disse que tem uma irmã... Você tem outros irmãos?”) Quantos anos eles têm?
10. Seus pais são daqui de São Paulo mesmo? (Se sim, perguntar também sobre avós, bisavós... até encontrar a primeira geração da família que veio pra cá). Quantos anos eles (pais) têm? Quando seus (pais/avós/bisavós) vieram pra São Paulo? Você sabe por que eles vieram?
11. Você tem filhos? Quantos anos eles têm?
12. E o resto da família, tios, primos, também vivem aqui em São Paulo? (Se sim, em que bairros? Moram próximo? Se não, onde? Tem contato sempre?)

TRABALHO/OCUPAÇÃO (aprox. 5 min.)

Objetivos: obter informações sobre rede social do informante; características socioeconômicas

13. Você trabalha aqui por perto? (se não souber onde a pessoa trabalha)
14. O que você faz?
15. Você gosta do seu serviço? (Se não, o que a pessoa preferiria fazer?)

LAZER (aprox. 5 min.)

Objetivos: obter informações sobre rede social do informante; mobilidade na cidade; características socioeconômicas

16. E nas horas de lazer, o que você e sua família gostam de fazer? (Se saem, vão pra que lugares?) Você acha que a cidade de São Paulo tem boas opções de lazer? Quais?
17. A maioria dos seus amigos mora aqui nesse bairro mesmo? (Se não, onde?)
18. Você tem algum atividade de recreação em grupo, algum clube...?

19. Quais são seus amigos mais antigos? Você mantém contato com os amigos de colégio?
20. Você costuma viajar? Pra que lugares já viajou? Que lugares gostaria de conhecer?

II. Segunda parte A CIDADE DE SÃO PAULO (aprox. 20 min.)

21. Uma pesquisa recente mostrou que 57% das pessoas em São Paulo deixariam a cidade caso pudessem. O que você acha disso? (se o informante perguntar, a pesquisa é do Ibope e foi publicada em janeiro/2010 3)
22. Você gosta de morar em São Paulo? (Se não, em que lugar preferiria morar?) Por quê?
23. O que você acha que caracteriza a cidade (tanto as coisas boas quanto ruins)?
24. O que você acha que caracteriza o paulistano (tanto as coisas boas quanto ruins)?
25. Olhando pra mim, você diria que eu sou paulistano(a)? Por quê?
26. O que você mais gosta em São Paulo?
27. O que você não gosta em São Paulo? (a depender do tópico mencionado pelo informante - violência, trânsito, poluição etc. - procurar explorar mais o assunto. P.ex.: você já foi assaltado? O que aconteceu? (para obter narrativa pessoal) O que o governo poderia fazer pra solucionar esse problema? (para obter uma fala mais distanciada)) (explorar o subtópico por aproximadamente 10 min.)
28. Pras pessoas que não vivem em São Paulo, como você acha que elas imaginam que seja a cidade? Qual é a imagem que as pessoas de fora de São Paulo têm da cidade?
29. Você falou que já foi pra (X, Y, Z). Quando você foi pra esses lugares, as pessoas percebiam que você era paulistano? (Se sim) como elas percebiam?
30. Quando você conhece alguém, você percebe se a pessoa não é daqui de São Paulo?
- a. (Se sim) como você percebe? (Se o informante mencionar o modo de falar, seguir nessa linha e tentar conseguir informações mais precisas. É normal as pessoas não saberem definir o porquê, mas devemos tentar tirar mais informações).
- b. (Se não) Quando você ouve uma pessoa falando, por exemplo, você percebe que ela é de fora pelo sotaque?

PRODUÇÃO/PERCEPÇÃO/AVALIAÇÃO LINGÜÍSTICA (aprox. 10 min.)

31. Já que a gente está falando de sotaque... qual o sotaque do Brasil que você mais gosta? E tem algum que te irrita? Como é que (o gaúcho/o carioca/o mineiro/o caipira etc. – a depender dos sotaques mencionados) fala(m)?

32. Aqui em São Paulo tem muito (i)migrante, né? De onde vêm a maioria das pessoas? Tem algum bairro específico em que eles (italianos, coreanos, nordestinos etc.) se concentram?

33. E como é que as pessoas falam na cidade de São Paulo? (evitar usar a palavra “sotaque”)

34. E aqui dentro da cidade de São Paulo, você vê diferenças entre as diferentes regiões, diferentes bairros? (Se sim) Você poderia dar alguns exemplos? (É normal as pessoas responderem que não, mas se responderem que sim, tente obter informações mais precisas)

(Imprimir a lista de palavras, a notícia e o trecho, todos abaixo, para mostrar ao informante)

LISTA de PALAVRAS: Apresentar ao informante a lista de palavras e dizer: “Agora eu queria te pedir pra ler algumas coisas. Eu tenho aqui uma lista de palavras, e eu queria que você lesse cada uma delas. Você pode dar uma olhada primeiro, antes de começar a ler.” Esperar que o informante leia a lista de palavras. Depois perguntar:

35. A gente estava falando de sotaques agora há pouco... Como você acha que uma pessoa do interior falaria algumas dessas palavras? Tem diferenças?

36. E como um carioca falaria algumas dessas palavras?

37. Tem mais algum sotaque no Brasil que você conhece?

LEITURA DE NOTÍCIA: Dizer ao informante: “Agora eu queria que você lesse esse texto. Você pode ler em voz baixa antes de falar em voz alta.” Esperar que o informante leia a notícia.

LEITURA DE TRECHO: Dizer ao informante: “Por último, queria que você lesse esse trequinho.” Esperar que o informante leia o trecho.

38. O que você acha desse modo de falar: “Você tá entendendo o que eu tô dizendo?” (com “en” ditongado e exagerado). Se a pessoa manifestar uma atitude negativa, perguntar: Como você acha que deveria ser? O que tem de errado aí?

39. Quem você acha que fala assim? (se falar “paulistanos”, perguntar: “você acha que todos os paulistanos falam assim ou é uma coisa de uma região mais específica na cidade?”)

40. Você fala desse modo?

41. E o que você acha de “Me vê dois pastel e um chopps?” (Fazer as mesmas perguntas acima)

42. E o que você acha de “A porta tá aberta.” (com retroflexo exagerado). Fazer as mesmas perguntas acima.

43. E tem mais algum modo de falar que você acha assim bem paulistano?

Lista de palavras

alma	erguer	orca
amargo	fazenda	ordem e progresso
animação	felicidade	orgânico
argola	firme	órgão
atitude	fornalha	ostracismo
barqueiro	furgão	penteadado
biscoito	fusquinha	pertencimento
cacique	geleira	perto
carteiro	gérmen de trigo	porto
cerca	gordo	presente
chácara	gula	rapidez
circo	Hércules	riqueza
cisne	Hilário	sabor de menta
curto	Irmã	soberba
defender	Justiça	trabalho
discoteca	Lento	turco
elefante	mortadela Marba,	urgente
enchente	mosca	utilidade
entender	necessidade	vulto
entretenimento	noite	zebra

Notícia

08/02/2010

Moradores cobram solução para área alagada em SP e aguardam encontro com prefeito

Moradores da região do Jardim Pantanal protestaram nesta segunda-feira em frente à Prefeitura de São Paulo contra a inundação em bairros da zona leste, que completa dois meses

hoje. Uma comissão de manifestantes foi recebida pela Secretaria de Relações Institucionais e apresentou reivindicações, mas o prefeito Gilberto Kassab (DEM) não participou do encontro.

Representantes dos manifestantes esperam que o governador José Serra (PSDB) também participe da reunião de sexta. De acordo com os integrantes da comissão, não havia integrantes do governo estadual no encontro realizado hoje na prefeitura.

Além de cobrar a limpeza imediata das águas nos bairros, os moradores cobram uma solução para a falta de moradia. A maioria dos desalojados continua em escolas municipais e, com o início do ano letivo, temem não ter para onde ir.

Trecho

Tá chovendo muito! Choveu tanto, tanto na semana passada que ficou uma piscina na minha casa. Ó, pra você ver: Molharam todos os armários, a cama, os colchões, tudo... Foi um sacrifício... O que a gente fez? Nós tivemos que erguer os móveis pra limpar tudo: a geladeira, o forno... minha irmã até veio me ajudar, sabe? E meus filhos compraram umas cadeiras novas, mas é aquela coisa, assim... quando chover de novo, vai molhar tudo outra vez. Você fica sem ter o que fazer. E tem um rio lá perto que sempre alaga... quer dizer, é água dentro e fora de casa! Daí, o que acontece? Fica aquele trânsito, os carros todos parados, a gente demora um tempão pra chegar em casa... Não aguento mais enchente nessa cidade... Agora que eu vou fazer? Os políticos falam, falam, mas eles tinham que fazer alguma coisa urgente. Você tá entendendo o que eu tô dizendo?