

UM ESTUDO SOCIOLINGUÍSTICO DA CONCORDÂNCIA NOMINAL DE NÚMERO EM UMA COMUNIDADE DE CAXIAS - MA A SOCIOLINGUISTIC STUDY OF NOMINAL NUMBER AGREEMENT IN A COMMUNITY OF CAXIAS – MA

Catarina de Sena Sirqueira Mendes da Costa¹

Virna Pereira Teixeira²

Universidade Federal do Piauí

RESUMO

Nesta pesquisa, cujo *corpus* é constituído por entrevistas com 20 informantes, agrupados de acordo com variáveis sociais como escolaridade, faixa etária e sexo, são descritos e interpretados os registros de concordância nominal de número da fala de moradores de um bairro periférico de Caxias - MA. Com base na Sociolinguística variacionista (WEINREICH *et al.*, 2006; LABOV, 2008) e trabalhos sobre a concordância nominal de número no Português Brasileiro (SCHERRE, 1988; NARO e SCHERRE, 2007; LUCCHESI, 2009) foram analisados, com o auxílio do programa Goldvarb X as ocorrências de variação nos sintagmas nominais. Por se tratar de uma comunidade marcada por fluxos migratórios do campo para a cidade, os estudos sobre a influência desta dinâmica espacial para incorporação ou não dos usos mais prestigiados do português de Bortoni-Ricardo (2011) fazem parte também das reflexões deste trabalho. Em relação às variáveis linguísticas, os dados foram examinados segundo o princípio da saliência fônica (tonicidade do núcleo), classe morfológica do item pré-nuclear, posição linear e posição relativa dos constituintes. Identificou-se 1415 ocorrências da variante não prestigiada da concordância nominal de número, representando 62% dos sintagmas obtidos. Os resultados mostram ainda que os fatores mais atuantes para a realização da variante não prestigiada são a posição pré-nuclear do item constituinte (peso relativo correspondente a 0.95), nível médio de escolarização e 1ª posição do item constituinte (ambos com o mesmo peso relativo de 0.69). Sobre a caracterização da comunidade, mais de 80% dos sujeitos são naturais ou possuem vínculos familiares com a zona rural e têm contato regular com os meios de comunicação de massa. Logo, depreende-se que o uso da variante não prestigiada nesta comunidade é bastante expressivo, devendo ser considerada para compreensão deste fenômeno, sobretudo, a tendência de pluralização dos termos à esquerda do núcleo do sintagma e a atuação de um anterior vernáculo.

PALAVRAS-CHAVE: Variação linguística; Concordância nominal de número; Caxias.

ABSTRACT

In this research, whose *corpus* consists of interviews with 20 informants, grouped according to social variables such as schooling, age group and sex, the registers of nominal agreement of speech number of residents of a peripheral district of Caxias - MA are described and interpreted. Based on the variational sociolinguistics (WEINREICH *et al.*, 2006, LABOV, 2008) and papers on nominal number agreement in Brazilian Portuguese (SCHERRE, 1988; NARO and SCHERRE, 2007; LUCCHESI, 2009), the occurrences of variation in the noun phrases were analyzed using the Goldvarb X program. Since it is a community marked by migratory flows from the countryside to the city, the studies on the influence of this spatial dynamics to incorporate or not the most prestigious uses of Bortoni-Ricardo's Portuguese (2011) are also part

¹ Professora Doutora e Coordenadora do Mestrado em Letras da UFPI. E-mail: costacatarina@uol.com.br.

² Mestra em Letras pela Universidade Federal do Piauí. E-mail: virmateixeira@hotmail.com.

of the reflections of this work. Regarding the linguistic variables, the data were examined according to the principle of phonic salience (nucleus tone), morphological class of the pre-nuclear item, linear position and relative position of the constituents. It was identified 1415 occurrences of the non-prestigious variant of nominal number agreement, representing 62% of the obtained phrases. The results also show that the most active factors for the realization of the non-prestigious variant are the pre-nuclear position of the constituent item (relative weight corresponding to 0.95), mean level of schooling, and the first position of the constituent item (both with the same relative weight of 0.69). About the characterization of the community, more than 80% of the individuals are natural or have family ties to the rural area and they also have regular contact with the mass media. Therefore, it is possible to conclude that the use of the non-prestigious variant in this community is very significant, and should be considered in order to understand this phenomenon, especially, the tendency of pluralization of the terms to the left of the nucleus and the performance of a previous vernacular.

KEYWORDS: Linguistic variation; Nominal number agreement; Caxias.

INTRODUÇÃO

A variação da concordância nominal de número no português falado no Brasil é objeto amplamente pesquisado no país desde a década de 1970 (SCHERRE, 1988). Estigmatizada, a variante popular da concordância nominal de número (ausência parcial/total das marcas de plural nos constituintes do sintagma nominal) é recorrentemente associada a baixos níveis de escolarização. Segundo Bortoni-Ricardo (2011), esse fenômeno deve, ainda, ser observado de acordo com os deslocamentos de grandes contingentes populacionais que migraram da zona rural de diversas regiões brasileiras para as cidades, os quais passaram então a assimilar padrões urbanos, especialmente linguísticos, a fim de se adaptarem à nova realidade social.

Inicialmente, fez-se necessário investigar o fenômeno da variação da concordância nominal de número no município de Caxias-MA, tendo em vista a escassez de trabalhos que tratem deste objeto tomando por base a fala de seus moradores. A escolha da comunidade, o bairro Campo de Belém, destarte, diz respeito à caracterização social e geográfica da localidade. Embora situada na periferia da cidade, sua origem é predominantemente rural, o que permitiria compreendê-la também à luz das investigações sociolinguísticas de Bortoni-Ricardo (2011). O *continuum* rural-urbano, assim por ela definido, pôde ser melhor evidenciado com base no estudo das redes sociais dos falantes, proposta que será discutida neste trabalho como suporte para balizar algumas considerações sobre a realidade observada.

Acerca das teorias que buscam explicar a gênese da variação popular do Português Brasileiro (doravante PB), é importante registrar que Bortoni-Ricardo (2014) cita, em favor de um eventual processo de *pidginização* que teria ocorrido no Brasil, um fragmento de português falado por índios em 1620, trazido por Serafim da Silva Neto (*apud* BORTONI-RICARDO, 2014). Tais contatos linguísticos, neste período teriam, conforme a autora (2014, p. 32), influenciado principalmente os usos linguísticos das camadas populares da zona rural, uma vez que a população residente no campo à época era composta majoritariamente de negros e indígenas, algo que teria, portanto, contribuído fortemente para o emprego de variedades mais distantes das normas de prestígio do português europeu.

Por outro lado, as análises de Scherre e Naro (2007) tomam a hipótese da deriva secular, formulada por Edward Sapir (1980 [1949]), como ponto de partida para compreender o fenômeno da variação popular da concordância nominal e verbal do português falado no Brasil. Segundo esta hipótese, variações morfossintáticas do PB remontariam aos usos do português arcaico e, portanto, seriam extensões seculares de uma mesma variedade popular. Scherre (1988) focaliza, em particular, o papel das variáveis linguísticas para a realização da variante não

prestigiada da concordância nominal de número, as quais seriam bastante atuantes nesses usos em que a marcação de plural não é explícita em todos os itens constituintes dos sintagmas.

Com efeito, neste artigo são tratados inicialmente alguns tópicos sobre a constituição da Sociolinguística variacionista e seus fundamentos (WEINREICH *et al.*, 2006, LABOV, 2008), tendo como objetivo apreender o fenômeno da variação da concordância nominal de número no PB à luz de seus pressupostos. Em seguida são examinados, sucintamente, trabalhos relevantes da Sociolinguística no Brasil, com destaque para as publicações que realizam reflexões sobre a concordância nominal de número no PB. A partir de um *corpus* composto por 20 entrevistas de informantes, distribuídos por sexo (feminino e masculino), duas faixas etárias (18-34 anos e 35-50 anos) e dois níveis de escolaridade (ensino médio e superior), são verificadas as ocorrências de concordância de número nos sintagmas nominais segundo as variáveis sociais mencionadas e fatores linguísticos (saliência fônica - tonicidade do item nuclear, classe gramatical, posição linear e relativa), com o auxílio do programa Goldvarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005).

Os resultados desse levantamento quantitativo permitem aferir o grau de influência das variáveis linguísticas nas realizações da variante não prestigiada (ausência total/parcial de marcas de plural nos itens constituintes do sintagma nominal), assim como os grupos de fatores sociais que favorecem esta escolha dos falantes.

A coleta de informações sobre a naturalidade e origem dos familiares (rural/urbano), a abordagem dos hábitos religiosos dos moradores do bairro Campo de Belém, assim como os dados sobre contatos dos participantes com os meios de comunicação de massa, obtidos através das entrevistas, permitem compor um quadro interpretativo que busca apresentar evidências para um posterior estudo de redes, proposto por Bortoni-Ricardo (2011), tido como relevante nesta perspectiva, para tentar apreender o fenômeno da variação da concordância nominal de número através de categorias menos fixas.

Em síntese, é propósito deste trabalho analisar as ocorrências da variante não prestigiada levando em conta os resultados de estudos relevantes sobre este fenômeno do PB, os quais demonstram, especialmente, a importância do condicionamento dos fatores linguísticos nestas realizações. Além disso, busca-se cotejar outros os elementos sociais que precisam ser melhor apreendidos em investigações desta natureza, principalmente quando estão em perspectiva noções complexas como ruralidade/ urbanidade e ascensão social através da elevação dos níveis de escolaridade.

1 Fundamentos da Sociolinguística variacionista

No texto “Fundamentos empíricos para uma teoria da mudança linguística”, Weinreich, Labov e Herzog (2006 [1968]) discutem o modelo estruturalista, seja ele sincrônico, diacrônico ou ainda funcional, propondo fundamentalmente a incorporação da heterogeneidade da língua, a mudança, à teoria. A heterogeneidade passa então a ser o seu efetivo objeto de estudo, que deve ser analisada a partir de um tratamento sistemático.

Para que a mudança se estabilize em uma língua faz-se necessário, ainda, que a covariação de mudanças esteja associada ao longo do tempo e do espaço geográfico. Acrescente-se que as variações linguísticas e extralinguísticas estão inter-relacionadas, de forma que explicar a mudança linguística concentrando-se em apenas um dos dois aspectos não trará os resultados esperados para abarcar a complexidade da análise do comportamento linguístico através de sua observação empírica.

Outros sim, os autores elencam ainda cinco princípios empíricos, apresentados sob a forma de questionamentos, que norteariam a Teoria da Mudança linguística: o problema dos fatores condicionantes, o problema da transição, o problema do encaixamento na estrutura linguística e na estrutura social, o problema da avaliação e o problema da implementação (2006 [1968], p. 121-124). Como o recorte deste trabalho é sincrônico e não há estudo anterior sobre o objeto de estudo aqui selecionado na comunidade do bairro Campo de Belém ou mesmo no município de

Caxias, será colocado em perspectiva nas análises, especificamente, o problema dos fatores que condicionam a variação da concordância nominal de número nos sintagmas nesta localidade.

Dentre os autores desta obra fundamental para a expansão dos trabalhos sociolinguísticos, Labov notabilizou-se no meio acadêmico, sobretudo, por impulsionar ativamente as pesquisas da área nas mais diversas regiões do mundo e por empregar, com rigor metodológico, um tratamento quantitativo dos dados pesquisados. Ao investigar as variações linguísticas, através de métodos estatísticos, os estudos desenvolvidos por Labov sobre a mudança linguística ficaram conhecidos como Sociolinguística variacionista ou quantitativa.

Cabe registrar, preliminarmente, que o linguista norte-americano afirmava preocupar-se, sobretudo, com a formulação de uma metodologia consistente dos estudos sociolinguísticos, baseada no uso de testes, eliciações, intuições e observações do vernáculo, a qual fosse capaz de construir o caminho mais apropriado para a elaboração de uma teoria linguística (LABOV, 1972).

Entre as importantes contribuições da obra de Labov (1972) é importante destacar o conceito comunidade de fala. Segundo o autor (LABOV, 1972), trata-se de um conjunto de normas e atitudes compartilhadas por usuários de uma língua ou variedade linguística, a qual deve constituir objeto de estudos da Sociolinguística. Da mesma maneira, é a partir de uma determinada comunidade de fala que o investigador recolhe os dados para constituir um *corpus* e iniciar suas investigações.

Outra noção crucial nos trabalhos de Labov (2008[1972]) é o conceito de regra variável. Ao postular que toda variação é condicionada, ele define o conceito de regra variável em função de uma frequência significativa de usos linguísticos regidos por grupos de fatores (linguísticos e sociais). As formas linguísticas alternantes dá-se o nome de variantes, que por sua vez farão parte de um fenômeno tido como variável dependente, caso haja a preservação do significado deste uso e possibilidade de ocorrência no mesmo contexto (2008[1972]).

Com efeito, a tarefa do sociolinguista, nessa ótica, tem como objetivo descrever estatisticamente as realizações de variantes específicas em uma comunidade de fala, circunscritas ao fenômeno variável observado, e analisá-las de acordo com um grupo de fatores. Importa, ainda, estabelecer relações entre os achados obtidos em um tempo real, sincrônico e seu possível encaixe em um processo de mudança linguística que ocorre historicamente, sob um ponto de vista diacrônico.

2 Variação, fluxos migratórios e concordância nominal de número no português brasileiro

Levando em conta a complexidade da formação sócio-histórica brasileira e suas diferenças dialetais, Bortoni- Ricardo (2011) ao investigar os usos das variedades populares, tratou de elementos específicos da realidade nacional em suas análises sociolinguísticas do português falado no país. A autora afirma que o fenômeno de urbanização dos dialetos rurais estaria na base de todos os processos de mudança e estandartização do português no Brasil. Como uma das principais hipóteses levantada por Bortoni-Ricardo em seu trabalho, a autora faz o seguinte questionamento: “[...] quais são os principais fatores atuantes na manutenção de variedades rurais e/ou não padrão no Brasil?” (2011, p. 14).

Em substituição ao modelo homogêneo de análise, a autora opta por descrever e interpretar as variedades do português brasileiro ao longo de um *continuum* dialetal, que compreende desde vernáculos rurais isolados à variedade urbana de maior prestígio (2011, p. 22). Entre estes polos linguísticos estariam também variedades não padrão urbanas, utilizadas pelos segmentos mais pobres da população (analfabetos e semialfabetizados). Nestas, as marcas dos dialetos rurais ainda se fariam presentes, daí a autora denominar as variedades não padrão urbanas como variedades *rurbanas*.

Com efeito, a análise de redes, de acordo com Bortoni-Ricardo (2011), teria o propósito de compreender esse processo de transição do rural para o urbano, vislumbrando, sobretudo, como se dá a integração deste migrante na cultura urbana, que, ao se apropriar de diversos papéis

culturais, insere-se cada vez mais nestes ambientes complexos, ampliando o distanciamento em relação à sua comunidade de origem, tradicional e hegemônica. Daí a importância da distinção entre redes isoladas e redes integradas.

No que diz respeito às origens das variedades populares do PB, embora não haja uma defesa explícita em seus trabalhos da chamada hipótese crioulista, a autora admite que os contatos linguísticos entre indígenas, africanos e portugueses tenham favorecido a emergência de usos não prestigiados do português (BORTONI-RICARDO, 2014). Em outra linha de pesquisa, por sua vez, situam-se os trabalhos de Scherre (1988) e Naro e Scherre (2007), os quais se detêm mais objetivamente sobre a concordância nominal de número. Scherre, mais especificamente, em sua tese de doutorado, *Reanálise da Concordância Nominal em Português*, publicada em 1988, procurou descrever e explicar a variação da concordância nominal em uma língua específica, o português, ao tempo em que apresentou novos aspectos para serem tomados como princípios universais na teoria geral da linguística.

Nesta perspectiva de análise, marcas precedentes e posição linear dos constituintes do sintagma nominal levaram a autora a aferir que “marcas conduzem a marcas e zeros conduzem a zeros” (1988, p. 511), sendo que a unidade sintagmática composta por elementos nominais regulares e paroxítonos tende a ser menos pluralizada. Ao lado disso, quanto maior a diferenciação fônica entre as formas singular/plural, maior a probabilidade de marcação, especialmente quando sobre a sílaba tônica recair acento.

Os itens não nucleares em posição anterior ao núcleo do SN recebem mais marcação de plural, ao contrário dos itens pospostos, de modo que os núcleos de primeira posição apresentaram os resultados mais significativos quanto à formalização de plural. Ademais, Scherre (1988) argumenta que diante das evidências encontradas não foi possível comprovar a hipótese da descrioulização, uma vez que a autora teria verificado que as marcas de plural não ocorrem somente em posição inicial do SN e sim em qualquer outra posição, desvalidando, dessa maneira, o que seria um dos argumentos mais importantes da hipótese crioulista.

Por seu turno, tomando como premissa a noção da deriva secular proposta por Sapir (1980 [1949]), Naro e Scherre (2007) argumentam que seja no campo fonológico ou morfossintático as variações populares do PB já estariam previstas há séculos atrás no português falado em continente europeu.

Ao contrário do que propugnam Naro e Scherre (2007) acerca da hipótese de nativização do português, Lucchesi (2009) tratou da variação da concordância de número no português afro-brasileiro, originado, segundo o autor, nesta obra, a partir de uma transmissão linguística irregular³ do tipo leve. Ao pesquisar a fala dos moradores de Helvécia, distrito da zona rural do município de Nova Viçosa – BA, por exemplo, os resultados mostraram que apenas 10% da população realizavam a concordância nominal de número prestigiada.

Em comparação com outras variedades do português brasileiro, Lucchesi (2009) observou, igualmente, que a frequência de concordância entre estes termos, sujeito e predicativo/ sujeito e estruturas passivas, é relativamente alta entre os falantes da norma urbana semiculta (ao redor de 50%), entretanto ela se mostrou praticamente inexistente nas comunidades rurais afro-brasileiras isoladas. Por fim, o autor reconhece que o contato dessas comunidades com as formas de prestígio urbanas estaria modificando os usos linguísticos destes falantes em favor da norma culta, processo este liderado essencialmente por homens, jovens e que viveram certo período fora de suas localidades.

³ Segundo Lucchesi (2009, p. 510-511), ao contrário dos países colonizados pelos portugueses em que houve a formação de crioulos típicos, no Brasil, o contato linguístico entre línguas de origem africana e a língua portuguesa não constituiu propriamente crioulos em território nacional, mas a partir dessa convivência linguística o português passou a adquirir características crioulistas, ainda presentes nas variedades rurais do PB, daí o autor afirmar que este processo deve ser tratado como uma transmissão linguística irregular do tipo leve.

Considerando, portanto, as reflexões sobre a Sociolinguística variacionista (WEINREICH; LABOV; HERZOG, 2006 [1968]) e seu refinamento metodológico desenvolvido por LABOV (2008[1972]), as análises sobre o continuum rural/urbano de Bortoni-Ricardo (2011) e os estudos de Scherre (1988) e Naro e Scherre (2007), são apresentadas adiante as estratégias de investigação empregadas e os resultados obtidos com base na amostra selecionada.

3. O corpus

O município de Caxias localiza-se na região do Leste maranhense, a 360 quilômetros de distância da capital do estado, São Luís. Atualmente, dos 155.202 habitantes de Caxias, segundo dados do Censo do IBGE de 2010, 118.559 vivem na zona urbana e 36.643 na zona rural, o que significa dizer que um pouco mais de 75% da população municipal reside no perímetro urbano do município.

A localidade em estudo, o bairro Campo de Belém, surge, por sua vez, na década de 1970 a partir de um projeto habitacional que contou com a parceria entre a Igreja Católica do município, inicialmente proprietária de todo o perímetro em que se situa atualmente o bairro, e a Prefeitura de Caxias, tendo como propósito assegurar moradia às pessoas de baixa renda vindas, sobretudo, da zona rural do município.

Sobre a amostra, em relação à variável sexo, dos 20 informantes, 11 são mulheres e os outros 9 são homens. Decidiu-se também estabelecer apenas duas faixas etárias para as análises: o primeiro segmento, que se convencionará chamar de Faixa I, compreende os falantes que possuem entre 18 a 34 anos, representando o grupo de jovens adultos da pesquisa; o segundo segmento, Faixa II, foram incluídos os participantes que têm entre 35 a 50 anos de idade, identificados como adultos maduros. Em relação à escolarização, agruparam-se os sujeitos da pesquisa em dois níveis, assim denominados: Nível I de escolaridade, formado por aqueles que possuem o Ensino Médio completo ou estão em fase de conclusão e Nível II de escolaridade, constituído por indivíduos com ensino superior completo ou graduandos.

4. Metodologia

O foco da pesquisa realizada é uma variável bastante estigmatizada (ex.: "Os *menino*" / "as outras *pessoa*") no Brasil e por haver divergências entre os linguistas do que de fato seria a norma padrão da língua portuguesa em uso no país (FARACO, 2008), com base em Camacho (2013) adota-se nesta investigação a terminologia variante prestigiada/ não prestigiada ou desprestigiada para referir-se aos usos linguísticos identificados na amostra.

Como dito anteriormente, a variável dependente analisada é a realização da concordância nominal de número em sintagmas nominais. Como só são possíveis duas formas alternantes, ausência parcial/total ou marcação canônica de plural em todos os elementos do sintagma nominal, para aferir as probabilidades de maior ocorrência de registros em que não há a marcação de plural em pelo menos um dos constituintes, fez-se necessário o levantamento de todos os sintagmas nominais presentes nas falas dos informantes.

Em relação às variáveis independentes, optou-se por trabalhar com três variáveis sociais (sexo, faixa etária e nível de escolarização), as quais permitem estabelecer relações com outras pesquisas da área. Ao lado destas, foram elencadas como variáveis linguísticas, neste estudo, a posição relativa e linear dos itens constituintes do sintagma nominal (analisam-se os constituintes antepostos ou pospostos aos núcleos do sintagma, sua posição linear e classe gramatical), assim como a saliência fônica (tonicidade das sílabas) do núcleo do sintagma nominal.

Partindo de uma análise atomística, com o suporte quantitativo do Goldvarb X, foram incluídos 2 participantes em cada célula. Esta distribuição habilitou, portanto, a construção de 8 células, nas quais há o acréscimo de um terceiro informante em 4 células para totalizar o conjunto de 20 indivíduos.

Sem necessariamente figurarem como variáveis, através das entrevistas foi possível obter também dados relativos à origem/ naturalidade dos informantes (nascimento ou parentesco com

moradores da zona rural do município, região ou estado), contato com os meios de comunicação de massa (televisão, Internet, revistas e jornais impressos) e hábitos religiosos.

5. Análise dos resultados

Foram identificadas, ao todo, 2.274 ocorrências de concordância nominal de número no português falado da comunidade do bairro Campo de Belém. Destas, 1.415 são ocorrências da variante não prestigiada e 859 são ocorrências da sua variante prestigiada. É necessário destacar que praticamente 90% da amostra é constituída por sintagmas nominais simples (dois constituintes, sejam eles determinantes, quantificadores, adjetivos ligados aos nomes ou ainda categorias substantivadas). O processamento estatístico do programa revelou que todas as rodadas foram estatisticamente significativas, de modo que são dispostos a seguir os números conforme os grupos de fatores sociais e mais adiante de acordo com as variáveis linguísticas selecionadas.

5.1 Grupos de fatores sociais

Após contabilizadas as ocorrências e a rotação dos dados, o percentual relativo à quantidade de sintagmas da variante não prestigiada, mostrou-se bastante superior em relação às ocorrências da variante prestigiada, em todos os grupos de fatores sociais. A seguir são apresentados os valores para aferir a significância de motivadores externos nestas realizações:

Tabela 01 - Análise dos grupos de fatores sociais - variante não prestigiada

GRUPO DE FATORES	FATORES	TOTAL	%	PESO RELATIVO
SEXO	f	610/973	62,7	0.50
	m	805/1301	61,9	0.49
<i>Significance 0.694</i>				
FAIXA ETÁRIA	a	744/1111	67	0.5
	j	671/1163	57,7	0.45
<i>Significance 0.000</i>				
ESCOLARIDADE	s	581/756	76,9	0.66
	t	834/1518	54,9	0.41
<i>Significance 0.000</i>				

Fonte: Elaboração da autora

No grupo de fatores sexo, os valores de pesos relativos para informantes do sexo feminino e masculino são praticamente equivalentes, embora, em termos percentuais, entre as informantes do sexo feminino haja um número ligeiramente superior de ocorrências da variante não prestigiada.

Em relação à faixa etária, os pesos calculados também estão próximos, 0.55 para informantes adultos e 0.45 para informantes jovens, contudo a diferença é pouco mais relevante no caso de informantes adultos, os quais foram responsáveis pela maior quantidade de ocorrências da variante não prestigiada neste grupo de fatores.

Os pesos relativos mais expressivos encontram-se no grupo de fatores nível de escolarização. Indiscutivelmente, o segmento de informantes com formação básica, cujo peso é de 0.661, figura como o fator que mais influencia a realização da variante não prestigiada, em oposição ao fator que compreende informantes com formação superior. É importante destacar que o acesso dos moradores deste bairro ao ensino superior é bastante recente. A expansão das vagas nas universidades na última década em todo o país possibilitou a estes moradores a

participação em um universo acadêmico antes intransponível para as comunidades periféricas, de maneira que a ampliação dos anos de escolarização permitiu, nesta localidade, maior contato e assimilação da variante prestigiada da concordância nominal de número.

5.2 Variáveis linguísticas

a) Saliência fônica - Tonicidade do item nuclear

Em princípio, para investigar a influência da variável linguística saliência fônica, como ocorre com frequência em trabalhos dessa natureza, foram analisados os usos linguísticos verificando se as formas mais salientes de plural são mais marcadas que as menos salientes. Há dois eixos para o estudo dessa variável: grau de diferenciação material fônica na relação singular/plural e tonicidade da sílaba.

Ocorre que, após a codificação da amostra, majoritariamente os núcleos dos sintagmas levantados apresentavam pluralização regular (a pessoa/ as pessoas) e raros eram os casos de plurais irregulares (o mês/ os meses, a oração/ as orações) ou ainda metafônicos (o posto/os postos). Nesse sentido, considerando, preliminarmente, que este fator linguístico teria influência limitada, os núcleos do sintagma foram analisados sob o ponto de vista de uma escala baseada na tonicidade das sílabas (oxítonas singulares e monossílabos tônicos provavelmente mais marcados que paroxítonas e proparoxítonas), e examinado o acento deste item constituinte em sua forma singular. Assim sendo, a Tabela 02 traz um levantamento de todos os sintagmas identificados, segundo a tonicidade dos núcleos (sílabas oxítonas, paroxítonas e proparoxítonas), como pode ser visualizado adiante:

Tabela 2 - Saliência fônica - tonicidade do núcleo singular/ variante não prestigiada

GRUPO DE FATORES	FATORES	TOTAL	%	PESO RELATIVO
Saliência fônica Tonicidade do Item Singular	Oxítona Ex.: essas <i>questão</i> (mjt, 31 anos)	284/475	59,8	0.47
	Paroxítona Ex.: nas <i>porta</i> (fjt, 18 anos)	1107/1768	62,6	0.5
	Proparoxítona Ex.: as <i>característica</i> (mat, 38 anos)	24/31	77,4	0.67
				<i>Significance 0.101</i>

Fonte: Elaboração da autora

Os núcleos com sílabas oxítonas foram os que menos influenciaram a realização da variante não prestigiada. Os núcleos com paroxítonas não se mostraram influentes, já as proparoxítonas são determinantes, nessa amostra, para a não realização da variante prestigiada. De acordo com Scherre (1988, p. 54), as marcas de plural seriam mais frequentes em monossílabos tônicos ou em oxítonas singulares, ao passo que nos itens lexicais paroxítonos e proparoxítonos haveria uma tendência de não marcação explícita de plural, uma vez que o morfema de plural não é encaixado, em língua portuguesa, nas sílabas paroxítonas e proparoxítonas. Com efeito, neste *corpus*, embora numericamente pouco expressivos, os itens lexicais nucleares proparoxítonos foram os mais significativos para a realização da variante não formal.

b) Posição linear dos itens constituintes

Não se pode considerar isoladamente a posição linear dos itens constituintes para inferir a atuação de fatores linguísticos na realização da concordância nominal de número. É necessário

relacioná-la a outros fatores, como classe gramatical dos itens constituintes. Embora, como já foi afirmado, os sintagmas nominais nesta amostra sejam preponderantemente simples, com os itens ocupando as 1ª e 2ª posições nas ocorrências, a inclusão deste fator contribui para verificar a significância das 3ª e 4ª posições nestes usos. Vale dizer que a 5ª posição foi excluída por figurar como *knockout* na primeira rodagem. Seguem os resultados do levantamento dos sintagmas nominais, sob a análise deste fator:

Tabela 03 - Posição linear dos constituintes/ variante não prestigiada

GRUPO DE FATORES	FATORES	TOTAL	%	PESO RELATIVO
Posição linear dos constituintes	1ª Posição Ex.: os <i>posto</i> (fas, 35 anos)	1193/1240	96,2	0.88
	2ª Posição Ex.: <i>pras casa albeia</i> (fjs, 18 anos)	180/871	20,7	0,07
	3ª Posição Ex.: nos pequenos <i>comércio</i> (mas, 37 anos)	35/126	27,8	0.1
	4ª Posição Ex.: as minhas duas <i>vizinha</i> (fjt, 18 anos)	7/37	18,9	0.06
<i>Significance 0.000</i>				

Fonte: Elaboração da autora.

Os valores probabilísticos atribuídos as 2ª, 3ª e 4ª posição foram muito baixos e isto permite afirmar que estes fatores linguísticos não são significantes, nesta investigação, para a realização da variante desprestigiada. Ao contrário, o item constituinte que ocupa a 1ª posição no sintagma nominal é recorrentemente pluralizado, mas as ocorrências nas quais ele foi observado não há realização da concordância formal, de modo que o peso correspondente a 0.887 é bastante consistente para admitir a importância deste fator nos usos da variante não prestigiada.

c) Classe Morfológica do item plural/pluralizável pré-nuclear

Considerando que, nesta amostra, os núcleos dos sintagmas nominais são, predominantemente, pertencentes à classe morfológica dos substantivos, foram definidos como itens para a investigação deste fator linguístico somente os termos que precediam imediatamente o núcleo dos sintagmas. Apesar de a classe dos numerais não flexionar nos casos de concordância observados, ela foi incluída, como item pluralizável, por remeter à ideia de quantidade, de elementos variados, similar, portanto, ao conteúdo do item plural. Seguem, na Tabela 04, os dados probabilísticos de todas as classes analisadas para exame dos fatores linguísticos mais influentes no uso da variante não prestigiada:

Tabela 04 - Classe morfológica do item plural/pluralizável pré-nuclear/ variante não prestigiada

GRUPO DE FATORES	FATORES	TOTAL	%	PESO RELATIVO
Classe gramatical do item plural/pluralizável pré-nuclear	Numeral Ex.: vinte <i>ano</i> /vinte anos (fjt, 27 anos/ mat, 38 anos)	277/517	53,6	0.41
	Categoria substantivada Ex.: os <i>outro</i> / os outros (fas, 35/ mjt, 27	10/22	45,5	0.33

	anos)			
	Artigo Ex.: uns <i>menino</i> / uns dias (fjs, 18 anos/ faz, 40 anos)	491/710	69,2	0.57
	Adjetivo Ex.: muitos <i>compromisso</i> grandes amigos (fat, 38 anos/ fjt, 26 anos)	74/130	56,9	0.44
	Pronome Ex.: meus <i>filho</i> / essas festas (fas, 38 anos/ mjs, 26 anos)	310/463	67	0.54
	Contração - preposição + artigo Ex.: das <i>criança</i> / das coisas boas (mjt, 31 anos/ fas, 44 anos)	161/273	59	0.46
	Quantificador Ex.: todas as disciplina/ todas as disciplinas (mjt, 27 anos)	37/64	57,8	0.45
<i>Significance 0.000</i>				

Fonte: Elaboração da autora

Os pesos atribuídos às classes morfológicas dos artigos e pronomes (determinantes), 0.57 e 0.54 respectivamente, foram os únicos que se mostraram atuantes nas ocorrências da variante não prestigiada. Logo, quando estes dois itens recebem marcação de plural há maior favorecimento para a realização da variante desprestigiada da concordância nominal. É importante observar que os valores dos pesos destas duas classes apontam apenas para uma ligeira significância, acima de 0.5, o que denota que estes fatores exerceram influência relativa quanto aos usos da variante da concordância não formal.

d) Posição relativa

Para efeito de melhor visualização dos resultados, foram assinaladas as posições dos itens constituintes imediatamente anteriores e posteriores ao núcleo dos sintagmas, denominadas respectivamente como pré-nuclear e pós-nuclear. A posição nuclear foi analisada principalmente para verificar a pluralização do próprio núcleo, mais recorrente nas ocorrências da variante prestigiada, tendo em vista que os casos de não marcação de plural do item pré-nuclear em 1ª posição foram excepcionais.

Feita esta consideração, encontram-se descritos, portanto, na Tabela 05, dados relativos à distância que os itens mantêm em relação ao núcleo e o papel que sua centralidade exerce:

Tabela 05 - Posição relativa/ variante não prestigiada

GRUPO DE FATORES	FATORES	TOTAL	%	PESO RELATIVO
Posição relativa	Nuclear Ex.: esses grandes pregadores/ <i>meu</i> tempos (mas, 42 anos/faz, 40	10/744	1,3	0.03

	anos)			
	Pré-nuclear Ex.: todos os bairros/ aqueles <i>riacho</i> (fat, 39 anos/ faz, 35 anos)	1237/1250	99	0.99
	Pós-nuclear Ex.: carros pesados/umas casinha bonitinha (mjt, 27 anos/ fat, 38 anos)	168/280	60	0.67
TOTAL				
				<i>Significance 0.000</i>

Fonte: Elaboração da autora

Com estes resultados, fica evidente que os termos pré-nucleares, lideram o favorecimento das marcas pluralizadas que se situam à esquerda do núcleo de sintagmas da variante não prestigiada. Já o núcleo com marca de plural, ocupando a 2ª posição como elemento constituinte nos sintagmas, ou na posição nuclear, como foi nesta pesquisa classificado este fator, é mais recorrente em sintagmas da variante prestigiada, por isso apresentando baixa influência na realização da concordância não formal. Os itens constituintes que sucedem o núcleo, por sua vez, também são atuantes para os usos da variante não prestigiada, com peso relativo de 0.67.

e) Reflexões sobre convergência e interação das variáveis sociais e linguísticas

Foram processadas, com os sete grupos de fatores em estudo (sexo, faixa etária, nível de escolaridade, tonicidade do item singular, posição linear, classe morfológica e posição relativa), 49 rodadas e selecionados os grupos de fatores mais atuantes na realização da variante não prestigiada e os grupos menos. Os grupos de fatores sexo, faixa Etária, tonicidade do item singular e classe morfológica foram avaliados como menos significantes para a realização da concordância não formal. Já os grupos de fatores considerados mais influentes, nível de escolaridade, posição linear e posição relativa, apresentaram os seguintes dados probabilísticos, dispostos na Tabela 06, logo abaixo:

Tabela 06 - Convergência entre os grupos de fatores mais influentes - variante não prestigiada

GRUPO DE FATORES	FATOR	PESO RELATIVO
Nível de Escolaridade	Nível I - Ensino médio	0.694
	Nível II - Ensino Superior	0.399
Posição Linear	1ª posição	0.696
	2ª posição	0.269
	3ª posição	0.365
	4ª posição	0.090
Posição Relativa	Pré- nuclear	0.952
	Nuclear	0.009
	Pós-nuclear	0.312
		<i>Significance 0.000</i>

Fonte: Elaboração da autora

O fator com maior significância é a Posição pré-nuclear do item constituinte dos sintagmas nominais. Como visto nos exemplos anteriores, a pluralização deste item, à esquerda do núcleo, desfavorece a realização da concordância nominal formal. Igualmente, a 1ª posição ocupada por itens constituintes influencia fortemente a realização da variante não prestigiada.

Dentre os fatores sociais, os informantes que possuem o ensino médio completo são os que mais realizam a variante não prestigiada, de maneira que este fator social mostrou-se o mais relevante nos usos linguísticos não formais observados nesta análise.

5.3 Dados relativos ao *continuum* rural-urbano

Os participantes desta investigação responderam questões relativas à naturalidade, relações familiares com moradores da zona rural, contato com os meios de comunicação e hábitos religiosos. Cada sujeito da pesquisa informou se era natural da zona urbana/rural de Caxias ou de outros municípios maranhenses (só foram admitidos como entrevistados nesta investigação participantes nascidos no município de Caxias ou em outros municípios do estado do Maranhão, desde que residentes em Caxias há mais de 2/3 de suas respectivas idades). Além disso, foram questionados se os seus parentes de 1º ou 2º grau eram oriundos da zona rural/urbana de Caxias, com que frequência mantinham contato com os meios de comunicação de massa e qual religião professavam. A síntese destes dados teve como objetivo proporcionar subsídios para uma reflexão mais qualitativa sobre os vínculos que a comunidade mantém com a zona rural do município, tendo como cenário o processo de urbanização que vivenciam no espaço da cidade. Na Tabela 07 foram consolidadas as informações ligadas às origens de cada entrevistado, conforme estão descritas a seguir:

Tabela 07 - Procedência dos informantes - rural/urbano

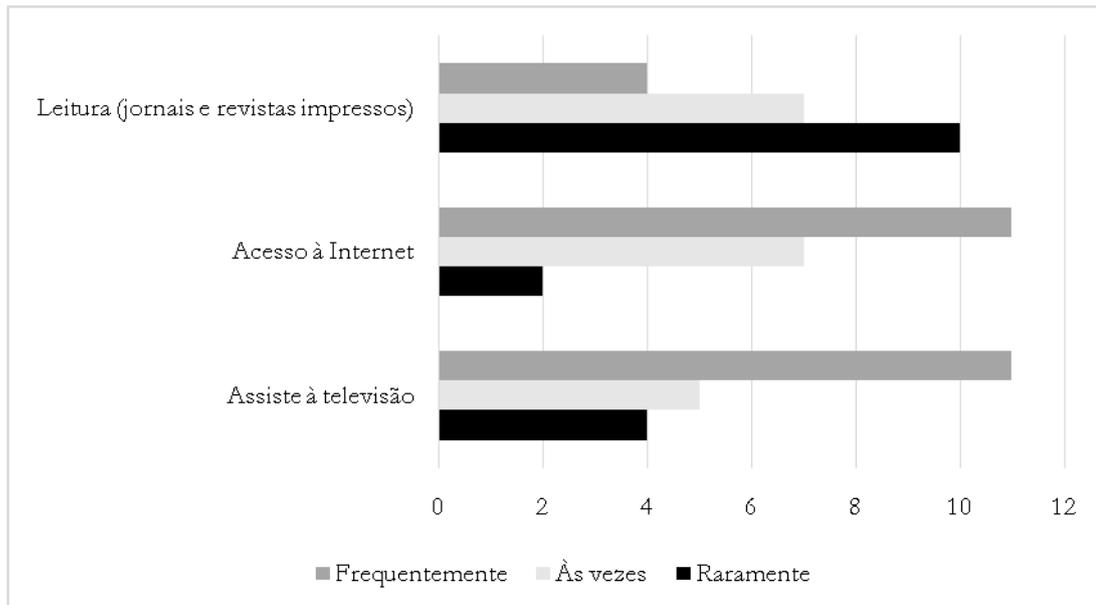
Naturalidade do informante	Número	%	Naturalidade dos familiares (pais/avós)	Número	%
Zona urbana de Caxias	9	45	Zona urbana de Caxias	1	5
Zona rural de Caxias	7	35	Zona rural de Caxias	16	80
Municípios da região	3	15	Municípios da região	3	15
Outros municípios do MA	1	5	Outros municípios	-	-
Total	20		Total	20	

Fonte: Elaboração da autora.

Os informantes quanto à naturalidade, em sua maioria, estão distribuídos entre a zona rural e urbana do município de Caxias, ou seja, 80% dos participantes. Em relação às origens dos familiares (pais, avós), 16 informantes, representando 80% dos sujeitos da pesquisa, afirmaram que seus parentes próximos nasceram e/ou ainda residem na zona rural do município.

No gráfico adiante, Figura 1, os resultados mostram o tipo de contato que os entrevistados têm com os meios de comunicação de massa (televisão, Internet, jornais e revistas impressos):

Figura 1 - Contato com os meios de comunicação



Fonte: Elaboração da autora

Somando os números de contato frequente e ocasional (às vezes), em ambos os meios de comunicação, Internet e televisão, cerca de 80% dos participantes estão familiarizados com estes veículos e são, portanto, alcançados pelos meios de comunicação de massa. Por outro lado, 50% dos entrevistados responderam que raramente leem jornais e ou/ revistas impressos, assim como outros 35% dos informantes, resultados que podem revelar tanto o hábito infrequente de leitura como a migração destes leitores para a Internet, que atualmente traz informativos e periódicos em plataforma digital.

Em relação aos hábitos religiosos, os entrevistados foram estimulados a discorrer sobre suas eventuais práticas religiosas, as quais evidenciarão os vínculos entre os moradores do bairro, a partir de redes sociais, ou ainda experiências de letramento promovidas por instituições religiosas. As informações obtidas estão consolidadas logo abaixo:

Tabela 08 - Hábitos religiosos

Respostas dos Informantes	Número	%
Católico praticante	10	50
Católico não praticante	7	35
Protestante	2	10
Cristão	1	5
Total	20	

Fonte: Elaboração da autora.

O grupo de pesquisados é constituído somente por adeptos de religiões cristãs. Há núcleos no bairro Campo de Belém formados por populações remanescentes das comunidades quilombolas rurais e que preservam os cultos das religiões de matrizes africanas, mas nenhum dos informantes se identificou como praticante destas religiões. A influência da Igreja Católica na edificação do bairro se faz presente nos hábitos religiosos de membros da comunidade. 85% do conjunto dos investigados são católicos, distribuídos entre católicos praticantes, representando 50% do total, e outros 35% que se declaram como não praticantes.

6 CONSIDERAÇÕES FINAIS

Os grupos de fatores sexo e faixa etária não foram significantes para a realização da variante desprestigiada. Isso importa dizer que a convivência com as marcas formais de concordância nominal é praticamente a mesma entre homens e mulheres, sejam eles mais jovens ou adultos maduros. Já o fator de escolarização nível médio, quando em convergência com os demais fatores, mostrou-se atuante nos usos da comunidade, com peso relativo de 0.69, o que permite considerar que o ingresso de sujeitos da pesquisa no ensino superior proporcionou a este grupo de informantes uma espécie de segunda etapa do letramento escolar, a qual sucederia a educação básica, e seria responsável por estimular mais o uso da variante prestigiada.

No grupo de fatores Posição linear, ficou demonstrado que dentre as quatro posições ocupadas, apenas a 1ª posição possui valor probabilístico significantes para a realização da variante não prestigiada, isto é, 0.95. Em outros termos, nos registros da variante não prestigiada a pluralização se aplica essencialmente aos itens que introduzem o sintagma nominal. De forma semelhante, no que diz respeito à Posição relativa, os itens que se destacaram na realização da concordância nominal não formal foram os termos que ocupam a posição pré-nuclear, enfatizando a posição à esquerda do núcleo como fator influente para a não marcação formal de plural.

Os participantes que migraram do campo para a cidade de Caxias, assim como seus familiares, 80% do universo de entrevistados, tiveram como objetivo, sobretudo, dar prosseguimento aos estudos e garantir melhores oportunidades de trabalho. Contudo, somente a escolarização de nível médio não garantiu a conquista de ocupações socialmente mais valorizadas e que requeiram o uso constante das variedades prestigiadas, de modo que os resultados significativos de ocorrências de concordância nominal não formal neste segmento estariam diretamente associados ao nível socioeconômico dos informantes e não somente à questão da escolaridade. O elevado contato com os meios de comunicação de massa também não se mostrou relevante para o favorecimento de usos prestigiados.

Sobre os hábitos religiosos, pode-se interpretar que o envolvimento dos informantes com práticas religiosas, principalmente da Igreja Católica, constitui importante fator de coesão social, especialmente em uma localidade onde "as pessoas se conhecem", como afirmaram os entrevistados. Trata-se, portanto, de uma comunidade *rurbana*, de passado recente, e que vive os desafios da transição para uma cultura urbana.

Sendo assim, a inclinação de pluralizar os primeiros itens do sintagma nominal, posicionados à esquerda do núcleo, há muito verificada em pesquisas pioneiras e nas mais atuais, mantém-se em meio urbano periférico, escolarizado. Mais pesquisas futuras serão então necessárias para aferir o encaixamento desta variação e sua homogeneização nos usos desta comunidade maranhense, que mesmo escolarizada ainda preserva em seus usos traços de um vernáculo rural.

[

REFERÊNCIAS BIBLIOGRÁFICAS

BORTONI-RICARDO, S. M. Do campo para a cidade: estudo sociolinguístico de migração e redes sociais. São Paulo: Parábola Editorial, 2011.

_____. Manual de Sociolinguística. São Paulo: Contexto, 2014.

CAMACHO, R.G. Da Linguística formal à Linguística social. São Paulo: Parábola, 2013.

FARACO, C.A. Norma culta brasileira: desatando alguns nós. São Paulo: Parábola Editorial, 2008.

IBGE. Instituto Brasileiro de Geografia e Estatística. Mapas interativos. Disponível em: <http://mapas.ibge.gov.br/interativos.html>. Acesso em jul.2016.

_____. Séries históricas e estatísticas. Disponível em: <http://seriesestatisticas.ibge.gov.br>. Acesso em 07 de janeiro de 2017.

LABOV, W. Padrões sociolinguísticos. São Paulo: Parábola Editorial, 2008 [1972].

- _____. Alguns princípios de metodologia linguística. *Linguagem em Sociedade*, v.1, abril de 1972, Universidade de Cambridge. p. 97-120.
- LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (orgs.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009.
- NARO, A.J.; SCHERRE, M. M. P. *Origens do português brasileiro*. São Paulo: Parábola Editorial, 2007.
- SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. *Goldvarb X: um aplicativo de regra variável para Macintosh e Windows*. Departamento de Linguística, Universidade de Toronto, 2005.
- SAPIR, E. *Linguagem: uma introdução do estudo da fala*. São Paulo: Editora Perspectiva, 1980[1949].
- SHERRE, M. M. P. *Reanálise da concordância nominal em português*. 1988. 554f. Tese (Tese de Doutorado em Linguística) – Universidade Federal de Rio de Janeiro, Rio de Janeiro.
- WEINREICH, U.; LABOV, W.; HERZOG, M.I. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola Editorial, 2006 [1968].

Recebido em 17/02/2017
Aceito em 14/06/2017
Publicado em 19/06/2017