



## VARIAÇÃO E O ESTATUTO DE VARIEDADES DO PORTUGUÊS

Silvia Figueiredo Brandão<sup>1</sup>

**RESUMO:** Focaliza-se, na perspectiva sociolinguística, à luz da Teoria da Variação e Mudança, a concordância de plural com base em estudos realizados pela autora, desde a década de 1990. Tais estudos abarcam a fala de comunidades pesqueiras do Rio de Janeiro e, ainda, variedades urbanas do Rio de Janeiro, de Lisboa e de São Tomé. As observações desenvolvem-se em duas etapas. Na primeira, centrada no Sintagma Nominal, parte-se da hipótese de que a análise comparativa do Português do Brasil (PB) com outras variedades não europeias é fundamental para que se possam melhor compreender as motivações que determinaram os padrões de concordância nelas registrados. Tecem-se, ainda, comentários sobre o papel do contato multilinguístico como elemento detonador dos padrões variáveis e discutem-se os conceitos de simplificação e complexificação linguísticas. Na segunda etapa, depois de mostrar o que se observa quanto à concordância nominal no Português Europeu (PE), centralizam-se as análises no PB e no PE, de modo a esboçar um quadro do estatuto das variedades do Português com base na tipologia de regras proposta por Labov (2003).

**PALAVRAS-CHAVE:** Concordância de número. Variedades do Português. Regras linguísticas. Multilinguismo.

**ABSTRACT:** Plural agreement is addressed from a sociolinguistic perspective, in the light of the Theory of Variation and Change and drawing on studies conducted by the author since the 1990s. These studies encompass the speech of fishing communities in Rio de Janeiro, as well as urban varieties in Rio de Janeiro, Lisbon and São Tomé. The observations were conducted in two stages. The first, centred on the Noun Phrase, starts from the hypothesis that analysis comparing Brazilian Portuguese (BP) with other non-European varieties is essential in order to understand better the motivations that have determined the patterns of agreement recorded in those varieties. Comments are also offered on the role of multilinguistic contact as a trigger for variable patterns and the concepts of linguistic simplification and complexification are discussed. In the second stage, after showing what is observed in noun agreement in European Portuguese (EP), the analyses centre on both BP and EP, to outline the status of the varieties of Portuguese based on the typology of rules proposed by Labov (2003).

---

<sup>1</sup> Professora Titular de Língua Portuguesa da Universidade Federal do Rio de Janeiro, Pesquisadora do CNPq e Cientista de nosso Estado, da FAPERJ. E-mail: [silvia.brandao@terra.com.br](mailto:silvia.brandao@terra.com.br).

**KEYWORDS:** Number agreement. Varieties of Portuguese. Linguistic rules. Multilingualism.

## Introdução

Nos primeiros estudos que se convencionou caracterizar como dialetológicos, alguns deles desenvolvidos por autodidatas em temas linguísticos, surgem as primeiras descrições de falares brasileiros, como é o caso do *Dialeto Caipira*, de Amadeu Amaral, publicado em 1920, que, além de um levantamento do léxico e de variantes fonéticas peculiares à fala do interior paulista, focaliza questões de concordância, tema que se mostraria bem mais complexo e abrangente do que a obra permitia, então, entrever.

Amaral observa, por exemplo, que a marca de plural se conserva nos constituintes “determinativos” do SN, mas se perde quando estes se tornam “pronominais” (“Estas carta não são as minha”). Também Nascentes (1922), ao descrever o dialeto carioca, chama a atenção para o fato de o morfema de número só permanecer nos determinativos, informação corroborada, ao tratar da variedade popular de Alagoas e Pernambuco, por Marroquim (1945), que, no entanto, adverte que os adjetivos, quando antepostos aos substantivos, apresentam a marca de plural (grandes coisa x coisas grande), já sinalizando a atuação da posição linear e relativa do constituinte no SN para a implementação ou não da regra de concordância nominal. Informações semelhantes são dadas, ainda, por José Ataliba Teixeira, que se ocupou dos dialetos mineiro (1983 [1938]) e goiano (1944). A concordância já mostrava ser um relevante fenômeno variável em diversos pontos do país.

Na obra desses precursores, encontra-se um repositório de algumas das características mais salientes do vernáculo do Português do Brasil (doravante PB), só bem mais tarde retomados de forma sistemática, graças, entre outros aspectos, à constituição, nas universidades brasileiras, de grupos de pesquisa de diferentes orientações teórico-metodológicas.

Não resta dúvida de que, no PB, a despeito da significativa variação que tem sido demonstrada em estudos que focalizam variáveis fonético-fonológicas, morfossintáticas e lexicais, entre outras, observa-se, do ponto de vista sintático ou morfossintático, a ocorrência de padrões similares em diferentes áreas do país, por exemplo, quanto à ordem dos clíticos, às estratégias de relativização, ao preenchimento do sujeito, ao objeto nulo.

No quadro da variação morfossintática, duas variáveis – a concordância nominal e a concordância verbal – se singularizam em relação ao quadro que acabo de esboçar, devido à sua valoração sociolinguística e à representatividade que lhe vem sendo atribuída no que tange aos debates sobre as origens do PB. É sobretudo por meio delas que, no PB e, como se verá adiante, também em países africanos, se pode traçar uma nítida fronteira entre variedades socialmente polarizadas, como as populares v. as cultas, as rurais v. as urbanas, esta última dicotomia aparentemente diatópica, mas, na verdade, social em essência, uma vez que as chamadas áreas rurais são constituídas, via de regra, por indivíduos que, além de não estarem inseridos em comunidades de grande estratificação social, tiveram pouco ou nenhum acesso aos meios formais de instrução. O nível de escolaridade, apesar das mudanças sociais que, nos últimos anos, se têm verificado na sociedade brasileira, pode, ainda, ser considerado o agente polarizador por excelência.

## Definição do tema

Nesta conferência, de forma sintética, trato da concordância de plural – um dos tópicos mais focalizados na linguística brasileira – na perspectiva sociolinguística, à luz da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968; LABOV, 1972, 1994, 2001, 2003). Complementarmente, apoio-me em artigo de Trudgill (2009a) que discute os conceitos de simplificação e complexificação linguísticas. Tenho como fontes principais de dados trabalhos que venho orientando e desenvolvendo, individualmente ou em parceria, desde a década de 1990, então no âmbito do Projeto do *Atlas Etnolinguístico dos Pescadores do Estado do Rio de Janeiro* (APERJ), e, nos últimos cinco anos, na esfera do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*.

## Caracterização dos corpora

O *corpus* APERJ constitui-se de entrevistas realizadas com 72 indivíduos (seis em cada localidade), todos pescadores artesanais, do sexo masculino, distribuídos por 13 comunidades do Norte e do Noroeste fluminenses e por três faixas etárias, com de zero a, no máximo, quatro anos de escolaridade quando entrevistados, em finais da década de 1980.

Já o *Corpus* Concordância, de perfil sociolinguístico bem diferente, retrata a fala de comunidades com alto grau de urbanização, com informantes (18 em cada comunidade) entrevistados entre 2008 e 2009 e distribuídos por sexo, três faixas etárias e três níveis de escolaridade. No que tange ao Português Europeu (doravante PE), ele abarca a fala de 54 indivíduos naturais de Cacém e Oeiras, na Região Metropolitana de Lisboa, e de Funchal, na Ilha da Madeira. No que concerne ao PB, são 36 indivíduos naturais de dois municípios da Região Metropolitana do Rio de Janeiro (o Município do Rio de Janeiro, representado pelos bairros de Copacabana/Botafogo, e pelo Município de Nova Iguaçu). Quanto ao Português de São Tomé (doravante PST), realizaram-se, ainda, análises com uma amostra alargada com 22 adultos de níveis fundamental e médio de instrução. Nove estudantes foram também objeto de investigação.

## Organização do texto e objetivos

Minhas observações desenvolvem-se em duas etapas. Na primeira, centrada no Sintagma Nominal, parto da hipótese de que a análise comparativa do Português do Brasil com outras variedades não europeias, aqui representadas sobretudo pela de São Tomé, é fundamental para que se possam melhor compreender as motivações que determinaram os padrões de concordância nelas registrados. Nesse sentido, comento duas variáveis em particular: (a) a *posição relativa e linear dos constituintes no SN*, que, dentre as de natureza estrutural, se vem mostrando a mais relevante para a determinação dos padrões variáveis de marcação de plural, e (b) o *nível de escolaridade* do falante, a mais saliente para a manutenção desses padrões e, por conseguinte, do quadro de polarização sociolinguística que se verifica no PB e no PST. Teço, ainda, comentários sobre o papel do contato multilinguístico como elemento detonador desses padrões variáveis, levando em conta, entre outras, as palavras de Petter (2007, p.9), que destaca a importância de se situar o PB “num conjunto maior – o dos países de fala portuguesa – onde se encontram outras variedades de português, não crioulas, faladas na África (Angola e Moçambique) e na

Ásia (Macau, Goa e Timor Leste)”, tendo em vista as “notáveis semelhanças que se observam, quanto à concordância de gênero e número no SN, nas variedades não-padrão do PB e nos crioulos de Guiné-Bissau, Cabo Verde, São Tomé e Príncipe”.

Na segunda etapa, depois de mostrar o que se observa quanto à concordância nominal no Português Europeu, centralizo minha análise no PB e no PE, com base em Labov (2003) para quem haveria três tipos de regras linguísticas: as *categóricas*, com 100% de aplicação, as *semicategóricas*, que ocorrem com 95% a 99% de frequência, e as *variáveis*, que têm de 5% a 95% de frequência (Cf. figura 1).

<b>Tipo de regra</b>	<b>Frequência com que opera</b>	<b>Violações</b>
Categórica	100%	Nenhuma, na fala natural
Semicategórica	95-99%	Rara e relatável
Variável	5-95%	Nenhuma por definição e não relatável

Fonte: Labov (2003)

**Figura 1.** Tipos de regras linguísticas.

Com apoio em resultados referentes a outras estruturas que implicam concordância, selecionados de pesquisas realizadas por Brandão (2011, 2013), por Brandão; Vieira (2012a, 2012b), Vieira; Bazenga (2013), Marcotulio; Vianna; Lopes (2013) e Vianna (2014), discuto o estatuto dessas duas variedades, partindo da hipótese de Vieira (2014) de que a tipologia de regras linguísticas proposta por Labov constituiria instrumento para propor perfis que caracterizariam o que costuma ser tratado, independentemente do quadro teórico adotado, como parâmetros gramaticais ou tipos linguísticos diversos.

## **A estrutura do SN**

Antes de comentar os padrões variáveis de marcação no SN, no PB e no PST, lembro o artigo em que Trudgill (2009a, p.98) sugere que “certos tipos de sociedade ou estrutura social tendem a produzir certos tipos de língua ou de estrutura linguística”. Segundo ele, dentre os fatores sociais mais propícios ao exame dos correlatos sociais da complexificação e da simplificação estruturais, estariam as dicotomias de base social “grau de contato *v.* isolamento”, “densidade *v.* debilidade das redes sociais” e “comunidades de pequeno *v.* grande porte”. Variedades linguísticas com uma história de alto nível de contato, sobretudo do tipo que envolve aprendizagem generalizada de uma língua por adultos, tenderiam a demonstrar altos níveis de simplificação, que estariam, por sua vez, vinculados aos processos de (i) regularização de irregularidades; (ii) aumento de transparência lexical e morfológica e (iii) perda de redundância.

Segundo ele, a perda de redundância assumiria duas formas principais: (a) a perda de categorias morfológicas, que pode ser compensada pelo uso de estruturas mais analíticas e (b) a redução na repetição de informação, como no caso da concordância gramatical.

Embora a simplificação e a complexificação possam ocorrer em qualquer língua e a simplificação

não se efetive só em situações de alto contato, é nestas últimas que há um desequilíbrio entre os dois processos, com o predomínio da simplificação (p.101). Nesse sentido, os pidgins representariam um caso extremo de simplificação e o “processo de expansão inerente à criouliização – o “reparo” da redução promovida pela pidginização – envolveria a reintrodução de algumas das categorias que foram perdidas” (p.103).

As observações de Trudgill parecem ir ao encontro do que se observa hoje em variedades não europeias do Português e do que provavelmente ocorreu no passado, sobretudo quando confrontadas à variedade que lhes deu origem, conforme se pode sustentar pelo Princípio do Uniformitarismo apresentado por Labov (1972) e segundo o qual processos que se observam no presente podem levar à depreensão de processos que ocorreram no passado.

No que se refere ao PB atual, a regra de indicação de plural sinaliza, sobretudo na variedade popular, um processo pretérito de simplificação e eliminação de redundâncias, enquanto o PE mantém a estrutura mais complexa, que implica a repetição da marca de plural em todos os constituintes flexionáveis do SN. Da mesma forma, como se verá adiante, em São Tomé, a variedade emergente do Português parece sofrer, ainda hoje, a influência das diversas línguas que coexistem na ilha, sobretudo do forro, ele também oriundo de um processo de simplificação.

### A variável posicional e questões relativas ao contato multilinguístico

Em todos os *corpora* referentes ao SN com que já tive a oportunidade de trabalhar, fica claro o predomínio, na modalidade oral, de SNs com núcleo na segunda posição, o que permite dizer que SNs de dois lugares são prototípicos da modalidade falada. No *Corpus* APERJ (cf. tabela 1), de um total de 2.865 SNs, 93,2% têm o núcleo na segunda posição. Nas amostras de Nova Iguaçu e de São Tomé, do *Corpus* Concordância, verifica-se a mesma tendência: 84%, em ambos os casos, o mesmo ocorrendo nas amostras do PE, com um índice ligeiramente menor: 79,4%.

Percentuais de SNs segundo a posição (P) do núcleo por variedade				
Corpus	Posição do vocábulo no SN			
	P1	P2	P3	P4
13 comunidades do N e NO do RJ	0,6%	93,2%	6,11%	0,03%
Nova Iguaçu	5,7%	84,2%	10,6%	9,4%
São Tomé	5,2%	84%	0,4%	0,5%

**Tabela 1:** Percentual de SNs segundo a posição (P) do núcleo em *corpora* referentes à comunidade de São Tomé e a comunidades brasileiras.

Em termos estruturais, portanto, ao se focalizar o SN na fala de comunidades quer pouco urbanizadas/rurais, quer altamente urbanizadas, lida-se com dados estruturalmente bem semelhantes (cf. tabela 2 e figura 2).

SNs VARIÁVEIS EM DOIS <i>CORPORA</i> REPRESENTATIVOS DO PB				
	SN 2	SN 3	SN 4	TOT.
SÃO TOMÉ	246 86%	37 13%	03 1%	286
NOVA IGUAÇU	255 82,5%	51 16,5%	03 1%	309
TOTAIS	501 84%	88 15%	06 1%	595

Nível	Copacabana	Nova Iguaçu
Fundamental	17	32
Médio	41	35
Superior	79	68

**Tabela 2:** Número de ocorrências e percentuais.

**Figura 2:** Índices percentuais.

Na fala dos indivíduos retratados no *Corpus* APERJ, o índice de indicação de plural nos constituintes do SN foi de apenas 44%, oscilando, nas diferentes comunidades, entre 38% e 53%. Para fins comparativos, lembro que Vieira (1995), na pesquisa sobre concordância verbal nas mesmas comunidades, obteve um total de 38% de aplicação da marca de P6, com oscilação entre 13% e 50%. (cf. tabela 3).

	Percentual geral de aplicação da marca de número		Oscilação na aplicação da marca de número nas comunidades	
	SN	P6 (Vieira)	SN	P6 (Vieira)
<b>Corpus APERJ</b>	44%	38%	13 comunidades 38% - 53%	13 comunidades 13% - 50%
<b>Corpus Concordância</b>	91,6%	83,5%	Rio de Janeiro 92,4%,	Rio de Janeiro 88,1%
			Nova Iguaçu 91,1%	Nova Iguaçu 78,2%

**Tabela 3:** Comparação dos percentuais de aplicação da marca de número no nome e na P6 do verbo em dois *corpora* representativos do PB.

Em ambos os casos, os falantes mais velhos apresentavam os maiores índices de não marcação de plural, sinalizando, como vêm demonstrando trabalhos mais recentes, uma tendência à aquisição e não à perda de marcas e permitindo estabelecer a hipótese de que, no PB, a forma conservadora ou vernacular seria a marcação defectiva de plural tanto em P6 quanto nos constituintes do SN.

Já no *Corpus* Concordância, observa-se, para o PB, um índice global de aplicação da marca da ordem de 91,6%, oscilando entre 92,4%, em Copacabana/Botafogo, e 91,1% em Nova Iguaçu. Vieira indica para a concordância verbal um índice global de 83,5%, sendo 88,1% em Copacabana e 78,2% em Nova Iguaçu. Quanto ao PST, a aplicação da marca é praticamente a mesma: 93,4% no SN e 93,1% em P6. Também nessas comunidades, em relação a ambas as variáveis, são os indivíduos mais velhos os que apresentam os menores índices de marcação.

Na variedade de São Tomé e em outras variedades africanas, empregadas como L1 ou L2, os padrões de indicação de número, como demonstram estudos sobre Moçambique, Angola, Cabo Verde (GONÇALVES 1997; INVERNO 2009; JON-AND 2010, 2011) são similares aos registrados no Português do Brasil (cf. Figura 3), cujas origens têm sido atribuídas por alguns pesquisadores (GUY 1981; HOLM 2002; LUCCHESI 2003, entre outros) ao contato multilinguístico que se verificou durante o processo de colonização.

Português de Angola (PA)	(a) quais são [os principais problemas]
	(b) [as coisa] estão muito cara
	(c) e buscar [as tuas fruta]
	(d) [os tais português], fomos [...]
Português de Moçambique (PM)	(a) graças a[os meus irmãos mais velhos]
	(b) tinha que levar [as criança]
	(c) [alguns nossos amigo] vê uma [...]
	(d) e d[os acontecimentos alegre]
	(e) boda de prata de [vinte e cinco ano]
Português de S. Tomé (PST)	(a) n[as antigas empresas coloniais]
	(b) morre [aquelas lula mesmo grande]
	(c) não era n[os nossos tempo]
	(d) pego n[as minha ferramenta]

Fontes: *Corpus Variedades Africanas do Português*, do Centro de Linguística da Universidade de Lisboa (PA e PM) e *Corpus VARPOR*, coletado por Tjerk Hagemeijer (PST).

Figura 3: Padrões de concordância nominal em três variedades africanas do Português.

As análises variacionistas realizadas com os mencionados *corpora*, em que a *posição relativa e linear do constituinte no SN* sempre foi selecionada como a primeira ou a segunda restrição mais importante para a presença/ausência da marca, permitem traçar um *continuum* de marcação de plural no SN, como o que propus em Brandão; Vieira (2012a), aqui acrescido dos índices referentes à análise com o *Corpus* APERJ. Tal *continuum* caracterizaria o PB e abarcaria, por hipótese, também as variedades africanas do Português, atendendo, assim, à proposta de Petter (2007) (cf. figura 4).

[+ marcas]				[- marcas]		
Pré-núcleo		Núcleo		Pós-núcleo		
Pos. 1	Pos. 2/3	Pos. 2	Pos. 3/4...	Pos. 2	Pos. 3	Pos. 4/5/6
.83 - .89	.50 - .69	.16 - .26	.16 - .32	.09 - .29	.10 - .16	.06 - .11
.97	.85	.12	.13		.15	

Figura 4: *Continuum* de marcação de plural no SN.

No pré-núcleo, a primeira e a segunda posições tendem a ser mais marcadas, o que sugere ser este o *locus*, por excelência, da marca. A partir do núcleo em segunda posição, a marca será menos frequente, em geral decrescendo a partir da terceira posição. No esquema proposto, a linha vertical contínua que separa o pré-núcleo do núcleo representa o limite da área de maior incidência de marcas, enquanto a linha pontilhada entre os constituintes nucleares e pós-nucleares indica que essas áreas não sofrem solução de continuidade quanto à não marcação.

No entanto, alguns estudos – como o de Baxter (2009) e o de Christino; Silva (2012) – sugerem outras configurações no que se refere, no primeiro caso, à área pré-nuclear e, no segundo, às áreas pré-nuclear e nuclear.

Baxter trabalhou com a fala de Helvécia, comunidade afrodescendente da Bahia, e com a comunidade dos Tongas, de São Tomé. Em ambas, os maiores índices de marcação de plural se concentram na posição pré-nuclear adjacente ao núcleo (como em *a minhas filha*). Os poucos constituintes à esquerda a ele não adjacentes apresentam menores índices de marcação, que, por sua vez, vão aumentando, gradualmente, da fala dos mais velhos para a dos mais novos. Levando-se em conta os índices globais por ele apresentados, verifica-se um *continuum* um tanto diferente daquele que propus, no que tange à área pré-nuclear. Baxter considera os constituintes à esquerda do núcleo elementos essenciais não só para a compreensão do mecanismo de concordância nessas comunidades, mas também para a aferição da gradativa aquisição do Português pelos indivíduos que as constituem, tecendo, ainda, uma série de considerações de cunho social para estabelecer convergências e divergências entre tais comunidades (cf. figura 5).

[+ marcas]		[- marcas]					
(-)	(+)						
Pré-núcleo		Núcleo		Pós-núcleo			
Pos. 1	Pos. 2/3	Pos. 2	Pos. 3/4...	Pos. 2	Pos. 3	Pos. 4/5/6	
.479	.914	.153	.397	<b>0% (90 ocos)</b>			<b>Helvécia</b>
.795	.915	.223	.267	<b>.112 (34/120 ocos)</b>			<b>Tongas</b>

Figura 5: *Continuum* de marcação de pluralidade em constituintes do SN duas comunidades afrodescendentes, elaborado com base em estudos de Baxter (2009: 278).

Christino; Silva (2012), por sua vez, focalizaram a concordância na modalidade escrita do Português Kaingang, com base em provas e trabalhos de professores indígenas em formação. Como nos informam as autoras, na língua Kaingang não há concordância, o que supostamente se reflete nas estruturas empregadas pelos indivíduos bilíngues (Português/Kaingang) quando se expressam em Português. Nos textos que lhes serviram de base, coexistem SNs inteiramente marcados quanto à concordância, SNs com marcação defectiva, típicos do português popular e, ainda, SNs sem marcação ou com marca apenas no núcleo, como no exemplo “e que cada aluno valorize [este curso], para sua vida profissional” e “trabalhar em [sua comunidades]” (p. 425). Quanto a este último padrão, empregado por 41,3% (57) dos autores dos textos, Christino; Silva (p. 425) acreditam que ele se deva à influência da língua Kaingang, “uma língua de cabeça à direita, em que a marcação das categorias morfossintáticas ocorre na porção direita do SN”.

Se levarmos em conta o princípio do Uniformitarismo, os estudos sobre a emergência, na atualidade, de variedades do português em comunidades indígenas podem também contribuir para a apreensão da dinâmica que deu origem ao PB e ao PST.

Nesse sentido, são importantes as observações de Emmerich; Paiva (2009: 162), quando focalizam o Português de contato, variedade utilizada pelos diferentes grupos étnico-linguísticos na região do Alto Xingu, em que o índice de concordância no SN é de 18%. Elas afirmam que essa taxa se deve ao fato de que, à medida que os índios vão ampliando seu contato com os falantes nativos do português, gradativamente se vai verificando um processo de despidginização, com a incorporação de padrões variáveis típicos do português popular como, no caso da indicação de plural no SN, a marcação nos constituintes mais à esquerda. As autoras mencionam, ainda, nesse particular, não só a atuação da variável posição linear em relação ao núcleo, mas também a do princípio da saliência fônica.

No âmbito do Projeto Concordância, realizei uma análise para testar se o uso de um crioulo, em especial do Forro, poderia estar interferindo no domínio da regra canônica de concordância.

A variedade de São Tomé pareceu-me adequada a testar essa hipótese, por diversos motivos, entre os quais, (a) contar-se com um *corpus* contemporâneo de fala urbana, representativo de indivíduos que se declararam falantes de português como L1; (b) ser essa a variedade, entre as dos países africanos que têm o português como língua oficial, que apresenta as características mais favoráveis a verificar tal interferência, pois, segundo Hagemer (2009), (i) o Português é falado por 98,9% da população, secundado pelo forro ou santomé, com 72,4% de falantes; (ii) lá coexistem, além do forro, outra língua crioula autóctone – o Angolar – e, ainda, o crioulo de Cabo Verde, o Lung'ie, da ilha do Príncipe, o Português dos Tongas e resquícios de línguas do grupo Banto; (iii) e São Tomé é o único país da África de língua portuguesa, que apresenta, “condições para a emergência de uma nova variedade” (p. 18). Com base nas respostas dos informantes quando o documentador lhes perguntava se falavam um crioulo, bem como em comentários e observações eventuais desses indivíduos sobre questões linguísticas, concluí haver, em São Tomé, indivíduos que (a) só usam o português e (i) não dominam ou (ii) dominam um crioulo; (b) usam o português e um ou mais crioulos (i) com menor ou (ii) maior frequência (no mercado, em momentos de lazer com amigos, com parentes mais velhos).

A variável *Frequência de uso de um crioulo* ficou, então, constituída por três fatores: frequência (a) *zero/baixa*, referente aos indivíduos que se expressam fundamentalmente em português; (b) *média*, relativa aos indivíduos que se expressam em português, mas dominam um crioulo e dele fazem uso eventualmente; (c) *alta*, abarcando os indivíduos que, embora falem o português e o tenham como L1, se expressam, regularmente, num crioulo (cf. tabela 4).

<b>Frequência de uso de um crioulo</b>	<b>Oco</b>	<b>Perc.</b>	<b>P.R.</b>
Zero /Baixa (grau zero - G0)	75/1039	7.2%	.41
Média (grau 1 - G1)	140/1117	12.5%	.52
Alta (grau 2 - G2)	90/219	41.1%	.75

**Tabela 4.** Atuação da variável *frequência de uso de um crioulo* para o cancelamento da marca de número no SN.

Independentemente da distribuição assimétrica dos informantes pelas três células, na análise variacionista sobre a fala de 22 indivíduos de níveis fundamental e médio, essa variável foi

selecionada como a terceira mais relevante, logo após, nesta ordem, a *posição linear e relativa do constituinte no SN* e o *nível de escolaridade*. Constatou-se que são os indivíduos que se comunicam preferencialmente num crioulo os que mais tendem a não implementar a marca de número (41.1%; p.r. .75). Com esses indivíduos, contrastam os que nunca ou pouco se expressam em crioulo (p.r. .41) e os que o utilizam apenas eventualmente (p.r. 52).

A questão da interferência de outras línguas sugerida pelos resultados ora apresentados há muito vem sendo debatida. A tendência à marcação de plural no constituinte pré-nuclear adjacente ao núcleo tem sido apontada, de acordo com Jon-And (2010, 2011), como argumento para uma possível influência de línguas africanas (transferência ou influência de substrato) em variedades não europeias do português. Segundo a autora, nos casos do Português de Moçambique e de São Tomé (o grupo dos Tongas e a comunidade de Almojarife), a hipótese se confirmaria, uma vez que essas variedades têm como substrato línguas Banto, em que a marcação de plural é feita por meio de prefixo.

Em síntese, os trabalhos aqui arrolados parecem indicar que variedades do Português oriundas de situações de contato linguístico revelam um nítido quadro de simplificação da expressão morfológica de pluralidade no SN.

### ***A variável Nível de escolaridade***

Nas análises variacionistas empreendidas no âmbito do Projeto Concordância, *nível de escolaridade* tem-se mostrado tão relevante quanto a *posição relativa e linear do constituinte no SN* para a (não)aplicação da marca de número, como, aliás, tem ficado patente em diversos outros estudos, inclusive aqueles relacionados a P6 (BRANDÃO; VIEIRA 2012a, 2012b; VIEIRA; BAZENGA 2013). A importância dessa variável reside no fato de, em geral, se associar a outras variáveis, como nível socioeconômico, maior ou menor acesso a bens culturais, maior ou menor contato com indivíduos de diferentes estratos sociais, maior ou menor consciência das variantes de prestígio. Ela é, portanto, um parâmetro fundamental para representar as distâncias sociais no Brasil – e também em São Tomé.

Surpreendem, sem dúvida, as elevadas taxas globais de marcação de plural no SN nas variedades urbanas dessas áreas (relembrando, 91,6% no PB e 93,4% no PST). No entanto, a atuação da variável *nível de escolaridade* mostra que tais índices mascaram a verdadeira realidade, uma realidade sociolinguisticamente polarizada, nos termos de Lucchesi (2003). Também obedecendo a um *continuum*, de forma escalar, os pesos relativos acentuam a distância entre os dois polos: os indivíduos de 5 a 8 anos de escolaridade e os de nível superior (cf. figura 6).

<b><i>Continuum de nível de escolaridade</i></b>				
	<b><i>[-marcas]</i></b>			<b><i>[+marcas]</i></b>
	<b><i>0 – 4 anos</i></b>	<b><i>5 – 8 anos</i></b>	<b><i>9 – 11 anos</i></b>	<b><i>12....</i></b>
NO e NE do RJ:	44%			
RMRJ-Copacabana		80,5% (.14)	93,7% (.37)	98,8% (.82)

RMRJ-Nova Iguaçu		91,4% (.38)	80,9% (.24)	97,2% (.76)
São Tomé		76,5% (.19)	96,4% (.49)	98,8% (.76)

NO e NE do RJ = Regiões Norte e Noroeste do Estado do Rio de Janeiro

RMRJ = Região Metropolitana da cidade do Rio de Janeiro

**Figura 6:** *Continuum* de marcação de número segundo o nível de escolaridade

Na amostra APERJ, não foi controlada a variável *nível de escolaridade*, uma vez que se lidava com um grupo homogêneo quanto a este particular, formado por indivíduos que ou não ou pouco frequentaram a escola: dos 72 informantes da pesquisa, 46 (64%) tinham de 1 até 4 anos de escolaridade e 26 (36%) eram analfabetos. Os pouco escolarizados e os analfabetos distribuíam-se pelas localidades de forma aleatória, no entanto, num *corpus* em que a aplicação da marca foi, como já se frisou, da ordem de 44%, em nenhuma das comunidades ela ultrapassou os 53%.

Embora não se tenha controlado sistematicamente a escolaridade, Almeida (1997) confrontou os analfabetos aos pouco escolarizados, deixando clara, de um lado, a homogeneidade do grupo, o que se verifica pelos índices percentuais bastante próximos, e mostrando, de outro, que, mesmo em comunidades tão semelhantes do ponto de vista sócio-econômico-cultural, a maior ou menor frequência, regular ou não, à escola pode ser fator determinante de diferenças, mínimas que sejam. Apenas 4% é o índice percentual de aplicação da marca que separa os analfabetos (42%) dos escolarizados (46%), embora a análise variacionista tenha indicado, respectivamente, pesos relativos .48 e .52, em decorrência da interação com as demais variáveis.

No que se refere a São Tomé, o cruzamento da variável *frequência de uso de um crioulo* com *nível de escolaridade* (cf. tabela 5) demonstra que são os indivíduos de nível fundamental os responsáveis pelos maiores índices de cancelamento da marca, que tem também caráter escalar: o cancelamento aumenta à medida que aumenta a utilização de um crioulo: baixa frequência: 19% < média: 25% < alta: 41%.

Frequência de uso de um crioulo	Nível Fundamental		Nível Médio	
	Oco	%	Oco	%
Baixa	52/276	19	23/763	3
Média	110/435	25	30/682	4
Alta	90/219	41		

**Tabela 5:** Cruzamento da variável Frequência de uso de um crioulo e nível de escolaridade. Aplicação: cancelamento da marca

Já os indivíduos de nível médio de escolaridade, com baixa ou com média frequência de uso de um crioulo, são os que menos cancelam a marca (respectivamente, 3% e 4%). O cruzamento desses dados, além de demonstrar que o domínio do padrão prestigiado de concordância pode estar sendo influenciado ou retardado pelo uso de uma língua em que a marca de plural não se aplica a todos os constituintes do SN, reforça a relevância da variável *nível de escolaridade*, que, de certa forma, sintetiza várias outras, de caráter sócio-econômico-cultural.

Apenas para complementar e finalizar esta seção, cabe dizer que, no PB e no PST, a marcação de plural no SN tem fortes condicionamentos estruturais, não se circunscrevendo a fatores referentes à saliência posicional aqui comentada, mas se relacionando ainda às saliências semântica e fônica: (i) o traço [+ animado] ou [-animado] do núcleo do SN repercute, respectivamente, na presença ou ausência da marca de número; (ii) menor ou maior grau de saliência fônica condiciona, inversamente, maior ou menor índice de marcação da pluralidade. No que concerne às variáveis sociais, também se observa a atuação de *Faixa etária*, no PB e de *Sexo*, no PB e no PST (Cf. tabela 6).

BP		STP
Copacabana	Nova Iguaçu	São Tomé
<u>NÍVEL DE ESCOLARIDADE</u>	<i>POSIÇÃO RELATIVA E LINEAR DO CONSTITUINTE NO SN</i>	<u>NÍVEL DE ESCOLARIDADE</u>
<i>POSIÇÃO RELATIVA E LINEAR DO CONSTITUINTE NO SN</i>	<u>NÍVEL DE ESCOLARIDADE</u>	<i>POSIÇÃO RELATIVA E LINEAR DO CONSTITUINTE NO SN</i>
<u>Faixa etária</u>	<u>Faixa etária</u>	<i>Animacidade do núcleo</i>
<u>Sexo</u>	Saliência fônica	<u>Sexo</u>
Saliência fônica	<i>Animacidade do núcleo</i>	Contexto subsequente
Significance: .000 Input: .98	Significance: .001 Input: .96	Significance: .023 Input: .98

**Tabela 6:** Variáveis atuantes para o cancelamento da marca de número no SN, por variedade.

A similaridade quanto à redução da expressão de marcas de plural (que parece registrar padrões diferenciados, provavelmente associados às línguas locais) bem como a atuação dessas variáveis estruturais nas diferentes variedades do Português sugere a existência de uma tendência de caráter universalizante, inerente às variedades do Português que emergem em áreas em que, no passado ou no presente, se registra contato com outras línguas, sobretudo as africanas, e, entre estas, crioulos de base portuguesa. Sugere, ainda, que tais padrões só se efetivam pela atuação de variáveis de natureza sócio-cultural, que *nível de escolaridade* parece sintetizar.

### A situação do PE

Em forte contraste com o quadro que se observa na variedade brasileira e na variedade santomense, no Português Europeu há efetivamente uma regra de concordância que determina que todos os constituintes flexionáveis do SN copiem a marca de plural presente no núcleo.

Como já ressaltai em Brandão (2013: 91),

nas gramáticas do século XVI e nas que as seguiram nos séculos XVII e XVIII, não há menção a padrões variáveis de concordância, embora diversas sejam as observações sobre aspectos fonético-fonológicos, em termos de correção e/ou de variação. Sabe-se que a variedade europeia passou por mudanças no século XVIII, mas estas parece não terem afetado o sistema flexional [...] Vale lembrar

que Peres & Moia (1995), que indicam a concordância como “uma das áreas críticas” da língua portuguesa, não arrolam casos variáveis na estrutura do SN.

No intuito de testar a hipótese de que, no Português Europeu, a regra de concordância nominal seria categoricamente implementada, organizei uma amostra formada por 6.952 SNs eliciados da fala de um total de 54 indivíduos (18 por localidade) nativos de Cacém e Oeiras, na Região Metropolitana de Lisboa, e de Funchal, na Ilha da Madeira.

Na amostra do Português Europeu (cf. tabela 7), como um todo, registraram-se 99,88% de aplicação da marca de número, não havendo entre as três comunidades nenhuma discrepância: os percentuais de aplicação da marca oscilaram entre 99,78% e 99,96%.

PE - Sintagmas nominais				
Amostra	Com marca(s) de número		Sem marca(s) de número	
	Nº de OCOs	%	Nº de OCOs	%
Oeiras	2310/2312	99,92	2/2312	0,08
Cacém	2448/2449	99,96	1/2449	0,04
Funchal	2186/2191	99,78	5/2191	0,22

Tabela 7: Distribuição dos dados referentes à concordância no PE.

Os 0,12% de seu cancelamento correspondem a 8 SNs (sete deles prototípicos de casos de marcação variável: dois constituintes, núcleo de baixa saliência fônica não marcado), dentre os quais seleciono dois para exemplificação: (a) a maioria d[as **peessoa**] **das mulheres** era tra/ era em casa era a bordar; (b) [**outros conhecimento**] **outra maneira** de ver as pessoas.

Os informantes, dois moradores de Funchal, ao enunciarem os primeiros dos SNs de cada exemplo, não aplicaram a marca no núcleo, mas imediatamente reformularam o enunciado por meio de outro SN – no primeiro caso, no plural; no segundo, no singular. Estes, como os outros seis casos, resultam, claramente, de falhas de processamento discursivo, de hesitação, ou lapsos, como assinalou Labov. Deve-se acrescentar que os oito SNs variáveis ocorreram na fala de sete diferentes indivíduos, no continente e na Ilha da Madeira, o que reforça tal interpretação.

Trudgill (2009), no artigo aqui já citado, afirma que, ao contrário do que ocorre na aprendizagem de uma língua por adultos em situação de contato multilinguístico, a aprendizagem de uma língua por crianças em contato coteritorial de longo prazo leva a uma complexidade adicional ou, eu diria, à manutenção de uma estrutura de maior complexidade, o que parece ser o caso do Português Europeu, pelo menos desde o século XVI, em que teve início a sua normatização.

## A concordância e o estatuto das variedades do português

Delineadas as características do PB e do PST, quanto a aspectos estruturais e sociais da aplicação da marca de número, contrastadas às do PE, de aplicação generalizada da regra, passo a comentar um estudo recentemente publicado (2014), em que Silvia Vieira e eu buscamos melhor compreender as diferenças que distanciam o Português do Brasil e o Português Europeu, analisando a concordância sob o prisma da tipologia das mencionadas regras linguísticas

propostas por Labov (2003). Para tanto, baseamo-nos não só nas análises quantitativas e qualitativas sobre a concordância nominal e verbal de P6 por nós realizadas, mas também, como já mencionado, em resultados de análises de outros pesquisadores referentes à concordância em estruturas com os pronomes *a gente* e *nós* (MARCOTULIO; VIANNA; LOPES, 2013; VIANNA, 2011), todas desenvolvidas no âmbito do Projeto binacional já citado.

Nesse texto, dizíamos que:

O registro de determinada estrutura em certo conjunto de dados e sua frequência de uso têm-se configurado como as evidências mais fortes para que o pesquisador a eleja ou não como opção prototípica de uma língua. Em pesquisas de natureza diacrônica, é exatamente o aumento gradativo da frequência de uso no decurso temporal que permite postular, ou ao menos sinalizar, um caminho de “mudança paramétrica”, “gramaticalização”, ou formulações semelhantes. Em outras palavras, classificar a regra linguística que envolve cada fenômeno – se categórica ou variável – permite não só estabelecer o parâmetro adotado em cada caso, mas também, no que tange à variação, avaliar sua extensão e seu perfil. Em se tratando de comportamento indubitavelmente categórico, tal estabelecimento é fácil e inquestionável. Quando se lida com comportamento variável, é a extensão dessa variabilidade que instaura o debate acerca do estatuto da língua/variedade. (p. 83).

Os resultados relativos ao Português do Brasil, no que tange à concordância nominal, já foram aqui comentados com base nas duas variáveis, uma linguística, outra social, que se mostraram mais salientes para a implementação ou não da marca de número. Quanto à concordância de P6, embora os índices percentuais globais de aplicação da marca sejam um pouco menores do que os registrados para a concordância nominal, observa-se a mesma nítida diferença entre as duas variedades.

Enquanto no PE (cf. tabela 8) a regra gira em torno dos 99% na Região Metropolitana de Lisboa e corresponda a 94,7% em Funchal, o que levou à caracterização da regra no PE como semicategórica, a regra no PB é francamente variável, com índices de 88,1% em Copacabana e 78,2% em Nova Iguaçu.

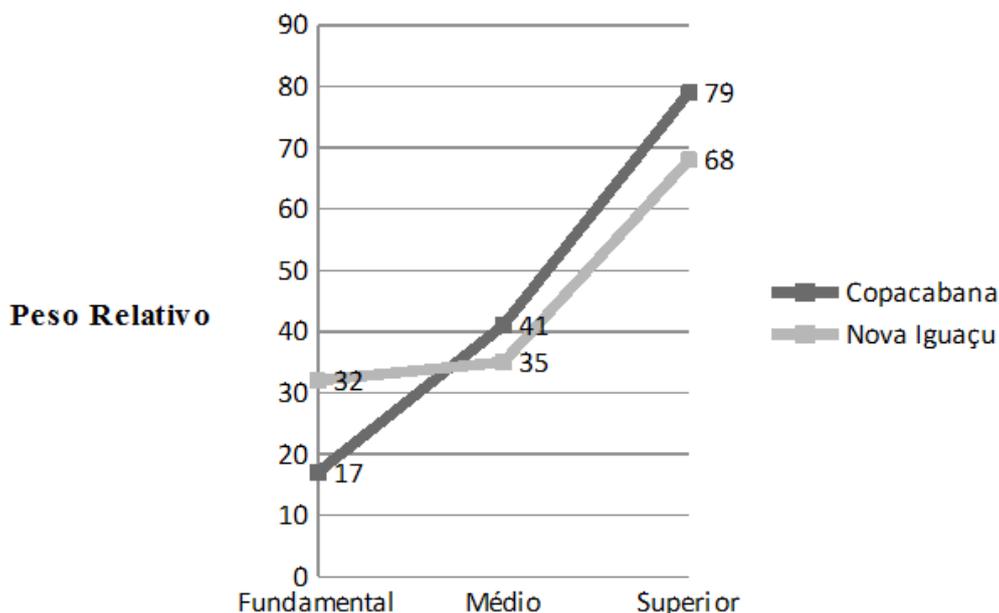
Amostra	Com marca(s) de número		Sem marca(s) de número	
	Nº de OCOs	%	Nº de OCOs	%
Oeiras	1454/1467	99,1	13/1467	0,9
Cacém	1176/1185	99,2	9/1185	0,8
Funchal	866/914	94,7	48/914	5,3
Copacabana	1229/1395	88,1	166/1395	11,9
Nova Iguaçu	1067/1365	78,2	298/1365	21,8

Fonte: Vieira; Bazenga (2013: 12)

**Tabela 8:** Distribuição dos dados com e sem marca de concordância verbal de P6 – PE e PB.

No âmbito de P6, no PB, as saliências posicional, fonética e semântica mostraram-se significativas para a interpretação do fenômeno, como têm demonstrado diversos outros estudos. O cancelamento da marca de número é favorecido em contextos com verbos cuja diferenciação entre singular e plural são de menor saliência fônica; com sujeito posposto; com o relativo “que”; com sujeito inanimado e verbos intransitivos, inacusativos ou copulativos. Tais contextos, que

se referem a estruturas que universalmente favoreceriam a não realização da marca, verificam-se “nas raríssimas ocorrências” da Região Metropolitana de Lisboa em que não se aplica a marca de número. Quanto a Funchal, além das mencionadas restrições, parece haver forte influência de condicionamentos fonético-fonológicos característicos da variedade madeirense, que sugerem a “necessidade de aprofundamento em relação aos chamados dialetos não-padrão do PE, tanto em sua caracterização linguística, quanto na descrição de sua sócio-história particular” (VIEIRA; BAZENGA, 2013). No que concerne à variável *nível de escolaridade*, que não se mostrou relevante no PE, ela tem caráter escalar e grande representatividade no PB (cf. figura 8).



Fonte: Vieira; Bazenga (2013: 25)

**Figura 8:** Concordância de P6 no PB: atuação da variável nível de escolaridade (em pesos relativos)

Quanto a construções com *nós* e com *a gente*, o PB e o PE também divergem (Cf. Tabela 9): quando a primeira pessoa do plural é expressa por *a gente*, no PB o verbo fica no singular, conforme a concordância canônica, em 99% das ocorrências, enquanto no PE, em 82% delas. Já quando se trata do pronome *nós*, o PE apresenta 100% de ocorrências no plural e o PB apresenta índices variáveis sobretudo em amostras de fala popular (cf. LUCCHESI; BAXTER; RIBEIRO, 2009), ou mesmo considerando estudos da fala urbana, conforme Vianna (2011), que registrou 90% de marcação de plural. Quando se trata de concordância nominal de número em construções predicativas com *a gente*, 21% delas, no PE, estão no singular, enquanto no PB, o singular é categórico (100%).

VARIEDADE		Conc. nominal de número (SN)	Conc. verbal (3ª p. pl.)	Conc. verbal (1ª p. pl.) <i>a gente</i>	Conc. verbal (1ª p. pl.) <i>nós</i>	Conc. nominal de número: construções predicativas com <i>a gente</i>
PE	Oeiras	99,9%	99,1%	82% (3SG)	100% (1PL)	21% (SG)
	Cacém	99,9%	99,2%			
	Funchal	99,7%	94,7%			

<b>PB</b>	Copacabana	92,4%	88,1%	99%	90%	100%
	Nova Iguaçu	91,1%	78,2%	(3SG)	(1PL)	(SG)

Fonte: Vieira; Brandão (2014)

**Tabela 9.** Índices percentuais referentes à concordância em diferentes estruturas.

Com base nos percentuais gerais obtidos dessa comparação (VIEIRA; BRANDÃO 2014: 103-107), chegou-se, preliminarmente, às seguintes conclusões resumidas na figura 9.

- (i) No que concerne à concordância nominal no âmbito do SN, PE e PB se opõem: no primeiro, a regra de marcação de pluralidade é categórica; no segundo, variável.
- (ii) Quanto à concordância verbal de 3ª pessoa plural, no PE continental a regra é categórica; em Funchal, está no limite entre o variável e o semicategórico; no PB, o comportamento é de uma regra variável. No caso de Funchal, a análise qualitativa dos dados – que demonstrou a especificidade de contextos de não realização da marca – permite categorizar o comportamento como semicategórico, à semelhança do ocorrido no PE continental.
- (iii) A concordância verbal de 1ª pessoa plural com *nós* revela uma regra categórica no PE e a não marcação de pluralidade (*nós trata*) é registrada exclusivamente no PB.
- (iv) No âmbito das estruturas de 1ª pessoa plural com *a gente*, a concordância verbal no PE registra não só a variante recomendada pela norma padrão (*a gente canta*), mas também a forma plural (*a gente cantamos*), em índices que permitem caracterizar a regra como variável, em contraposição ao PB, em que a opção pela forma singular é semicategórica.
- (v) No que se refere à concordância nominal em construções predicativas com *a gente*, o PB exibe uma regra categórica de realização singular e o PE, uma regra variável com preferência pela forma plural.

Estatuto da regra (Labov 2003)	Conc. no SN	Conc. verbal de 3ª p. pl.	Conc. verbal de 1ª p. pl. <i>a gente</i>	Conc. verbal de 1ª p. pl. <i>nós</i>	Conc. em predicativas com <i>a gente</i>
<b>Categórico</b>	<b>PE</b>			<b>PE</b>	<b>PB (singular)</b>

Semicatagórico		PE	PB (singular)		
Variável	PB	PB	PE	PB	PE

Fonte: Vieira; Brandão (2014).

**Figura 9:** Caracterização das diferentes regras.

Como salientamos (p. 103), a observação dos dados, ainda do ponto de vista quantitativo, permite estabelecer as seguintes generalizações:

- (i) O PB assume comportamento claramente variável em três das cinco estruturas controladas: concordância no âmbito do SN, concordância verbal de 3ª pessoa plural e concordância verbal de 1ª pessoa plural com *nós* (restrita, nesse caso, a variedades não prototipicamente urbanas e/ou escolarizadas). Em apenas uma das estruturas, exatamente a da concordância verbal com *a gente* (no singular), assume comportamento categórico; na outra estrutura com *a gente*, a nominal, exibe resultado semicategórico.
- (ii) O PE assume comportamento categórico / semicategórico em três das cinco estruturas controladas: concordância no âmbito do SN, concordância verbal de 3ª pessoa plural e concordância verbal de 1ª pessoa plural com *nós*. Apenas nas estruturas relacionadas a *a gente*, assume comportamento variável, visto registrar também a forma plural.

Mas o que se pode depreender dessas tendências, tendo em vista a hipótese que norteou a comparação, a de que a categorização proposta por Labov poderia servir de instrumento para a definição de tipos linguísticos?

Evidentemente que, apesar de ficar patente a sistemática diferença de comportamento do PE e do PB (não há coincidência quanto ao estatuto da regra em nenhum dos casos arrolados), é fundamental interpretar os dados também do ponto de vista qualitativo, o que aqui se tentará comentar resumidamente.

(a) O pronome *a gente*, de emprego mais recente, tem forma singular com valor semântico plural. Do ponto de vista canônico, fazer a concordância com *a gente* implica privilegiar a forma (verbo e predicativo no sing.), enquanto que com *nós* e *eles* ou com um SN no plural se privilegia a noção de plural, o que determina a aplicação da marca de plural.

O PE, quando não faz a concordância padrão, acaba optando por seguir o valor semântico, o plural, como em *a gente viemos*, em consonância com a opção pelo plural nos demais contextos. Já o PB, quando não faz a concordância canônica, opta pela forma de singular, em contextos formal e semanticamente de plural.

(b) No que respeita aos casos de concordância no SN e em P5 com o pronome *nós* (casos de concordância categórica no PE), fica claro que o *default*, nessa variedade, é o

plural, não tendo sido registradas, no caso de P5, ocorrências do tipo *nós trata*, como ocorre no PB.

(c) Quanto a P6, no PE, o conjunto de fatores em atuação parece obedecer a condicionamentos morfofonológicos (sândi externo) e sintático-semânticos do tipo genérico ou “universal” (cf. CHAMBERS 2004; TRUDGILL 2009; SZMRECSÁNYI; KORTMANN 2009), não podendo ser indicadores seguros do estatuto de cada uma das variedades do PE.

No PB, ao contrário, os dados encontrados exibem ausência de marca plural em frases com ordem direta não-marcada (sem qualquer conotação de estrutura topicalizada, por exemplo), com sujeito anteposto ao verbo, com formas verbais de alta saliência fônica e com traços animados, como no exemplo “*os menino foi na praia ontem*”.

Além de todas as evidências estruturais, ressalte-se a importância das restrições extralinguísticas no PB, dentre as quais se destaca o nível de escolaridade, sem nenhuma representatividade no PE, o que configura mais um índice de diferenciação entre essas variedades.

No PB, a regra básica, ou, em outros termos, vernacular, de concordância nominal, aqui exemplificada no sintagma “*os menino comportado*”, consiste na marcação do constituinte mais à esquerda. Sua repetição nos demais constituintes do SN está intrinsecamente vinculada à maior ou menor consciência do falante sobre o valor social da variável e ao maior ou menor domínio da regra canônica, em geral, aprendida na escola.

Em síntese, pode-se afirmar que, no PE, há uma efetiva regra de concordância de plural entre os constituintes do SN, enquanto no PB tem-se uma regra de marcação de plural que determina que essa categoria seja indicada no constituinte mais à esquerda (determinante ou qualquer outro elemento que possa ocupar a primeira posição).

O uso da regra de concordância nominal no PB seria uma decorrência de injunções de ordem social, uma regra artificial, resultante de alto grau de letramento, o que se verificaria também em variedades africanas, como a de São Tomé (cf. Brandão 2011, 2013), em que, em termos percentuais, se poderia definir a regra como semicatórica entre os indivíduos de nível superior. Isso justifica a constante menção, em estudos sobre o PB, ao aumento dos índices percentuais de concordância em função de escolaridade, embora os fatores que coatuem para a variação continuem a agir segundo os mesmos princípios, a mesma hierarquia. Nem na fala de indivíduos cultos a regra é categórica.

Vieira; Brandão (2014:108), nas considerações finais de seu artigo, afirmam acreditar que os resultados da análise da concordância por elas empreendida com base em parâmetros quantitativos e qualitativos constituam evidência de que a tipologia laboviana de regras linguísticas possa servir como instrumento para avaliar perfis tipológicos linguísticos em relação a qualquer fenômeno gramatical, sendo um forte recurso metodológico para promover a distinção entre o que é aparentemente variável e aquilo que constitui efetivamente uma tendência geral ou uma opção *default*. Sem dúvida, cumprir a agenda de pesquisa futura, por nós anunciada, permitirá testar a generalidade das nossas interpretações em amostras e temas diferentes.

## Considerações finais

Retomo uma de minhas observações iniciais, a de que a concordância de número é um dos tópicos mais focalizados na linguística brasileira. Tal observação poderia pressupor que pouco ainda haveria a dizer sobre o tema. No entanto, muitas são, ainda, as questões pendentes.

Para fornecer algumas das várias respostas que buscamos no sentido da caracterização do Português do Brasil e da recuperação de sua história, considero muito pertinente, além da sugerida utilização da tipologia de regras linguísticas, realizar análises contrastivas não só com o Português Europeu, mas também com variedades emergentes, sobretudo as de origem africana, que resultaram de situação de contato multilinguístico. Os padrões de marcação de plural no SN aqui indicados, por exemplo, justificam tal iniciativa. Nesse sentido, o melhor conhecimento também dos crioulos de base portuguesa e a verificação dos seus mecanismos de variação e mudança parecem ser mais dois relevantes tópicos para a recém mencionada agenda.

Lucchesi (2012: 45) caracteriza as comunidades afrodescendentes isoladas, em geral remanescentes de antigos quilombos (como a de Helvécia, por exemplo), como “sítios arqueológicos da história sociolinguística do Brasil”. Eu acrescentaria que, com o auxílio dos princípios e métodos da sociolinguística, considerar, qualitativa e quantitativamente, de forma mais detida, os crioulos de base portuguesa e as variedades emergentes do Português pode fornecer preciosas pistas para testar a hipótese da existência de um *continuum* afro-brasileiro do português, como sugerem direta ou indiretamente, entre outros, Petter (2007) e, ainda, Avelar; Galves (2014), em texto recentemente publicado.

Desse modo, espero que o conjunto de resultados e reflexões aqui apresentado possa, antes de trazer conclusões que se queiram definitivas para a caracterização do estatuto das variedades do Português como um todo, ter oferecido fortes evidências de que os estudos linguísticos se podem beneficiar fortemente da priorização de duas frentes de trabalho: (i) a descrição das variedades africanas do Português, seja na linha do que propõe Petter (2007), na definição de um continuum afro-brasileiro de aplicação das regras, seja na linha do que propõe Trudgill (2009), na observação dos processos de simplificação e complexificação na expressão linguística; e (ii) a interpretação das diversas regras linguísticas em questão segundo Labov (2003), acompanhada de adequada descrição quantitativa e qualitativa das ocorrências.

Amadeu Amaral – volto ao início deste texto e aos primórdios da descrição de falares do PB – em discurso na ABL, em 1925 (*apud* MENON 2008: 337), dizia, a propósito da comissão de brasileirismos que havia sido há pouco criada, que ela deveria ter “outra denominação e outro programa”, para que, ao invés de “só recolher e discutir palavras”, lançasse as bases

de um estudo metódico e puramente científico de todo o complexo de fenômenos dialectais no Brasil, fonéticos, morfológicos e sintáticos, investigando-lhes as causas e tendências, mas tratando antes de tudo de colhêr genuínos materiais de correcta observação, tomados diretamente à realidade viva da língua, fielmente transcritos segundo um sistema invariável, e não só se valendo de documento livresco ou dos depoimentos de segunda mão como de elementos subsidiários.

O que aqui apresentei é apenas um fragmento das pesquisas por mim desenvolvidas, por vezes com a colaboração de colegas e orientandos, a que se somam os estudos de vários outros pesquisadores também comprometidos com esse “programa”, tão antigo quanto atual, que se esboçou no início do século passado e começou a se efetivar a partir da década de 1970.

**Referências:**

- ALMEIDA, Evanilda Marins de. *A variação da concordância nominal num dialeto rural*. 1997. 145f. Dissertação (Mestrado em Língua Portuguesa). Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1997.
- AMARAL, Amadeu. *O dialeto caipira*. 3 ed. São Paulo: Hucitec; Secretaria de Cultura, Ciência e Tecnologia, 1976 [1920]. 197p.
- AVELAR, Juanito; GALVES, Charlotte. O papel das línguas africanas na emergência da gramática do português brasileiro. *Linguística* (Revista da ALFAL) 30 (2), pp. 241-288, 2014.
- BAXTER, Alan. A concordância de número. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.) *O Português afro-brasileiro*. Salvador, EDUFBA, 2009.
- BORTONI-RICARDO, Stella Maris. *Educação em língua materna: a sociolinguística na sala de aula*. São Paulo: Parábola Editorial, 2004.
- BRANDÃO, Sílvia Figueiredo. Em torno de um velho tema: o cancelamento da marca de número na fala das comunidades rurais brasileiras. *Revista Internacional de Língua Portuguesa* 12, pp. 50-57, dez. 1994.
- \_\_\_\_\_. O cancelamento da marca de número nominal na variedade urbana não standard do Português de São Tomé. XVI Congresso Internacional de la ALFAL, 2011, Alcalá de Henares. *Documentos para el Congreso Internacional de la ALFAL*. Alcalá de Henares: ALFAL/Universidad de Alcalá. Disponível em CD. Doc 104, 2011.
- \_\_\_\_\_. Patterns of plural agreement within the Noun Phrase. *Journal of Portuguese Linguistics* 12 (2), pp. 51-100, 2013.
- BRANDÃO, Sílvia Figueiredo; VIEIRA, Sílvia Rodrigues. A concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *Papia: Revista Brasileira de Estudos Crioulos e Similares* 22 (2), pp. 7-41, 2012a.
- \_\_\_\_\_. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa: Revista de Linguística* 56 (3), pp. 1035-1064, 2012b.
- CASTRO, Ivo. *Curso de história da língua portuguesa*. Lisboa: Universidade Aberta, 1991.
- CHAMBERS, Jack K. Dynamic typology and vernacular universals. In: KORTMANN, B. (ed.). *Dialectology meets typology. Dialect grammar from a cross-linguistic perspective*. Berlin/New York: Mouton de Gruyter, 2004. pp. 127-145.
- CHRISTINO, Beatriz; SILVA, Moana de Lima. Concordância verbal e nominal na escrita em Português-Kaingang. *Papia* 22 (2), pp. 415-428, 2012.
- EMMERICK, Charlotte; PAIVA, Maria da Conceição. Português xinguano: origem e trajetória. In: CARVALHO, Ana Maria. *Português em contato*. Madrid: Vervuert Verlag; Iberoamericana, 2009. pp. 153-163.
- GONÇALVES, Perpétua. Tipologia de “erros” do Português Oral de Maputo. In: STROUD, Christopher; GONÇALVES, Perpétua (org). *Panorama do Português Oral de Maputo: a construção de um banco de “erros”*. Moçambique: INDE, vol. II, 1997. pp. 37-70.
- GUY, Gregory Riordan. *Linguistic Variation in Brazilian portuguese: aspects of the Phonology*, *Diadorim*, Rio de Janeiro, Especial 2016, p. 83-104.

- Syntax, and Language History*. 1981. 391 fls. Ph.D. Dissertation (Linguistics). University of Pennsylvania, Pennsylvania, 1981.
- HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, Macau, 1 (1), pp. 1-27, 2009.
- HOLM, John. The social and linguistic origins of Brazilian Vernacular Portuguese: parallel cases of partial restructuring. *Revista Portuguesa de Filologia XXIV*, pp. 123-145, 2001-2002.
- INVERNO, Liliana. A transição de Angola para o português vernáculo: estudo morfossintático do sintagma nominal. In: CARVALHO, Ana (ed). *Português em contato*. Madrid: Veurvert Verlag; Iberoamericana, 2009. pp. 87-106. Disponível em: <[http://www.uc.pt/creolistics/research/angola/inverno\\_forthcoming](http://www.uc.pt/creolistics/research/angola/inverno_forthcoming)>. Acesso em: 05 mai. 2009.
- JON-AND, Anna. Concordância variável de número no SN no português L2 de Moçambique – algumas explicações sociais e linguísticas. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* 2, pp. 28-50, 2010.
- \_\_\_\_\_. *Variação, contato e mudança linguística em Moçambique e Cabo Verde. A concordância variável de número em sintagmas nominais do português*. 2011, 167 fls. Tese (Doutorado). Department of Spanish, Portuguese and Latin American Studies, Stockholm University, Stockholm, 2011.
- LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.
- \_\_\_\_\_. *Principles of linguistic change: internal factors*. Vol. 1, Oxford: Blackwell Publishers, 1994.
- \_\_\_\_\_. *Principles of linguistic change: social factors*. Vol. 2, Oxford: Blackwell Publishers. 2001.
- \_\_\_\_\_. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Org.). *Sociolinguistics: the essential readings*. Oxford: Blackwell, 2003. pp. 235-250.
- LEMLE, Miriam; NARO, Anthony Julius. Competências básicas do português. *Relatório final de pesquisa apresentado às instituições patrocinadoras Fundação Movimento Brasileiro de Alfabetização (MOBRAL) e Fundação Ford*. Rio de Janeiro, 1977. 151p.
- LUCCHESI, Dante. O conceito de transmissão linguística irregular e o processo de formação do Português do Brasil. In: RONCARATI, Cláudia e ABRAÇADO, Jussara (org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003. pp. 272-284.
- \_\_\_\_\_. A diferenciação da língua portuguesa no Brasil e o contato entre línguas. *Estudos de Linguística Galega* 4, pp. 45-65, 2012.
- \_\_\_\_\_; BAXTER, Alan; RIBEIRO, Ilza (orgs.) *O Português afro-brasileiro*. Salvador, EDUFBA, 2009.
- MARCOTULIO, Leonardo; VIANNA, Juliana; LOPES, Célia. Agreement patterns with ‘a gente’ in Portuguese. *Journal of Portuguese Linguistics* 12 (2), pp. 125-149, 2013.
- MARROQUIM, Mário *A língua do Nordeste* (Alagoas e Pernambuco). 2. ed. São Paulo: Companhia Editora Nacional, 1945.
- MENON, Odete Pereira da Silva. Uniformitarismo ou transmissão oral? In: VOTRE, Sebastião; RONCARATI, Cláudia. *Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica*. Rio de Janeiro: 7Letras, 2008. pp. 337-351.
- NARO, Anthony Julius; SCHERRE, Maria Marta Pereira. 2007. *Origens do português brasileiro*. São Paulo: Parábola Editorial, 2007.
- Diadorim*, Rio de Janeiro, Especial 2016, p. 83-104.

- NASCENTES, Antenor. *O linguajar carioca*. Rio de Janeiro: Organização Simões, 1953 [1922].
- PERES, João A.; MÓIA, Telmo. *Áreas críticas da língua portuguesa*. Lisboa: Caminho, 1995.
- PETTER, Margarida Maria Taddoni. Uma hipótese explicativa do contato entre o português e as línguas africanas. *Revista Papia* 17, pp. 9-19, 2007.
- SZMRECSANYI, Benedikt; KORTMANN, Bernd. "Between simplification and complexification: non-standard varieties of English around the world". In: SAMPSON, G; GIL, D.; TRUDGILL, P. (eds.). *Language complexity as evolving variable*. Oxford: OUP, 2009. pp. 64-79.
- TEIXEIRA, José Ataliba. *A linguagem de Goiás*. São Paulo: Anchieta, 1944.
- \_\_\_\_\_. O falar mineiro. Separata da *Revista do Arquivo Municipal* 45, São Paulo, 1938. 100 p.
- TEYSSIER, Paul. *História da língua portuguesa*. Tradução Celso Cunha. São Paulo: Martins Fontes, 1997.
- TRUDGILL, Peter. Sociolinguistic typology and complexification. In: SAMPSON, Geoffrey; GIL, David; TRUDGILL, Peter. *Language complexity as an evolving variable*. Oxford: University Press, 2009a. pp. 98-109.
- \_\_\_\_\_. Linguistic universals and vernacular data. In: FILPPULA, M.; KLEMOLA, J.; PITKÄNEN, H. (eds.). *Vernacular universals and language contacts*. New York/London: Routledge Taylor & Francis, 2009b. pp. 323-348.
- VIANNA, Juliana. *Semelhanças e diferenças na implementação de a gente em variedades do português*. 2011. 234 fls. Tese (Doutorado em Letras Vernáculas). Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.
- VIEIRA, Silvia Rodrigues. *Concordância verbal: variação em dialetos populares do Norte Fluminense*. 1995. Dissertação (Mestrado em Letras Vernáculas). Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1995.
- VIEIRA, Silvia Rodrigues. Entre o variável e o categórico: a concordância verbal e a colocação pronominal em variedades do Português. In: REZENDE, L. M.; SILVA, O. N. L. da; MENDONÇA, M. C.; ZAVAGLIA, C.; BRUNELLI, A. F. (org.). *A interdisciplinaridade e a especificidade linguística: teorias e práticas*. Araraquara; FCL-UNESP Laboratório Editorial; São Paulo: Cultura Acadêmica, 2014. (Série Trilhas linguísticas; 26) pp. 75-98.
- VIEIRA, Silvia Rodrigues; BRANDÃO, Silvia Figueiredo. Tipologia de regras linguísticas e estatuto das variedades/línguas: a concordância em português. *Linguística* 30 (2), pp. 81-112, dezembro, 2014.
- VIEIRA, Silvia Rodrigues; BAZENGA, Aline. Patterns of third person verbal agreement. *Journal of Portuguese Linguistics* 12 (2), pp. 7-50, 2013.
- WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical foundations for theory of linguistic change. In: LEHMANN, W.; MALKIEL, Y. (org.) *Directions for historical linguistics*. Austin: University of Texas Press, 1968. pp. 97-195.