
Concordância nominal em duas variedades do português: convergências e divergências

Silvia Figueiredo Brandão (UFRJ)

RESUMO: O artigo focaliza a concordância de número no sintagma nominal nas variedades urbanas do Português falado no Brasil e em São Tomé, com o objetivo de determinar os fatores que concorrem para a coexistência dos padrões variantes nelas encontrados. Os resultados, advindos de análise fundamentada nos pressupostos da Sociolinguística Variacionista, demonstram que há significativas convergências entre as duas variedades, entre elas as fortes restrições de ordem estrutural e social que norteiam o cancelamento do morfema de número.

Palavras-chave: concordância de número; sintagma nominal; português do Brasil; português de São Tomé; variedade urbana.

Introdução

Questões relativas à perda de morfologia flexional e de regras de concordância constituem importantes parâmetros para a definição do *status* de variedades que emergem de contato entre populações distintas dos pontos de vista linguístico e cultural. Nesse sentido, estudos sobre concordância verbal e nominal têm servido de base para a formulação de diferentes interpretações sobre o surgimento e desenvolvimento de variedades do Português.

Considerado um traço de diferenciação social de forte cunho estigmatizante e de caráter suprarregional, o cancelamento da marca de número no sintagma nominal (SN), e ainda no sintagma verbal, no Português do Brasil (doravante PB), vem sendo focalizado não só pelos pesquisadores interessados em determinar os condicionamentos sociolinguísticos dessas duas variáveis, mas também pelos que procuram caracterizar a variedade brasileira em oposição à europeia e, assim, esboçar a história do PB. Dentre as hipóteses apresentadas para a perda de morfologia flexional e de regras de concordância, destacam-se aqui a de Lucchesi (2003) e a de Naro; Scherre (2003a). O primeiro considera-a uma decorrência do processo de transmissão linguística irregular desencadeado pelo contato do Português com as diversas línguas africanas e indígenas que coexistiram no país a partir da fase de seu povoamento. Os últimos veem-na como o desenvolvimento de tendências existentes na deriva da língua de Portugal. Em outras palavras: de acordo com a primeira perspectiva, o processo de transmissão linguística irregular teria originado, no Brasil, novos processos de variação e

mudança, enquanto, de acordo com a última, o que ocorreu foi a ampliação de fenômenos já previstos no sistema.

Este estudo, norteado pelos pressupostos teórico-metodológicos da Sociolinguística Variacionista, tem caráter essencialmente descritivo e constitui uma abordagem preliminar, de cunho comparativo, sobre a concordância nominal de número na fala de comunidades urbanas do Brasil e de São Tomé, com o objetivo de verificar se os fatores que concorrem para a coexistência de padrões variantes no PB também atuam no Português de São Tomé (doravante PST), contribuindo, assim, para os debates sobre o processo de constituição das duas variedades.

A análise – que se apoia em *corpora* selecionados de 35 entrevistas (18 do PB, 17 do PST¹), com indivíduos distribuídos por sexo, faixa etária (18 a 35 anos; 36 a 55 anos; 56 a 75 anos) e nível de escolaridade (fundamental, médio e superior) – fundamentou-se no controle de quatro variáveis extralinguísticas e em variáveis estruturais que, no PB, se têm mostrado relevantes para a compreensão do fenômeno. Todos os informantes do PST são falantes de Português como Língua 1.

A amostra referente ao Brasil foi coletada em 2008-2009, em Nova Iguaçu, um dos vinte municípios que integram a Região Metropolitana do Rio de Janeiro, e a relativa a São Tomé foi selecionada de entrevistas realizadas, em 2008, por Tjerk Hagemeijer, pesquisador do Centro de Linguística da Universidade de Lisboa (CLUL).

1. O estudo da concordância nominal no PB: brevíssimo histórico, alguns resultados

Nos estudos pioneiros da dialetologia brasileira, já se teciam considerações, embora assistemáticas, sobre a concordância nominal e a simplificação do sistema de flexão de número, delineando-se, tenuemente, algumas das variáveis independentes que, mais tarde, se mostrariam importantes para a compreensão desses fenômenos. Para citar apenas um dialectólogo, Amaral (1976 [1920]), que focaliza o dialeto caipira de São Paulo, observa que a marca de plural se conserva nos constituintes “*determinativos*”, mas se perde quando estes se tornam “*pronominais*” (“*Estas carta não são as minha*”), acrescentando que o S e a vogal temática grafada *e* sofrem cancelamento em vocábulos terminados em consoante (os *ingrêis/os ingleses*; as *cor/as cores*), questão que, nos trabalhos da atualidade, veio a ser considerada sob o rótulo de saliência fônica.

Numa perspectiva dialectológica aliada a uma orientação gerativista, Rodrigues (1974: 68) retoma o dialeto caipira e afirma que, “na estrutura superficial, [...] os elementos redundantes de concordância são reduzidos de tal forma que há uma grande simplificação do sistema de concordância, com transformações morfofonológicas e morfosintáticas tanto no sintagma nominal sujeito como no sintagma verbal”. Mais tarde, Veado (1982), segundo as mesmas perspectiva e orientação teórica, abordando o dialeto rural de Minas Gerais, descreve a categoria de número tomando por base não só os elementos que estruturam o SN, mas também a posição superficial dos elementos do SN e sua relação com o Nome. A autora conclui que (a) constituintes antepostos ao Nome tendem a receber a marca (quando juntos, o artigo e o possessivo ocorrem no plural), enquanto os numerais inibem sua incidência; (b) constituintes que podem ocorrer antepostos ou pospostos ao Nome, no primeiro caso são marcados, no segundo, não, ressaltando que o pronome indefinido e o adjetivo pospostos não apresentam traço de número; (c) os Nomes só ocorrem com marca de número quando não precedidos de outro constituinte.

¹ Por não haver, entre os entrevistados, uma mulher de mais de 55 anos com nível superior de instrução, conta-se apenas com 17 informantes e não com 18, já que se considera um indivíduo por célula.

A partir de finais da década de 1970, surgem, na perspectiva sociolinguística variacionista, diversos trabalhos sobre o tema, alguns dos quais são aqui comentados à guisa de apresentação dos principais fatores condicionadores do fenômeno no Português do Brasil.

Braga; Scherre (1976) analisam a influência da posição do constituinte e da saliência fônica sobre a concordância no SN, concluindo que os constituintes mais sujeitos a receberem a marca de número são os que ocorrem na primeira posição e aqueles cujas formas singular e plural apresentam maior diferenciação fônica, o que é reiterado por outros estudos, como os de Braga (1977), Scherre (1978), Ponte (1979).

Guy (1981), em trabalho sobre o PB *substandard*, relaciona a variável classe gramatical à variável posição do constituinte no SN, tratando cada classe como privativa de uma determinada posição: o determinante, da primeira, o substantivo da segunda e o adjetivo, da terceira. Tal procedimento é criticado por Scherre (1988), que, apropriadamente, afirma que, pelo fato de tanto o substantivo quanto o adjetivo poderem ocorrer em outras posições, estabelecer correlações diretas entre classe e posição impediria a observação de qualquer regularidade que abranja outras classes.

Scherre, em sua tese de doutoramento (1988), procede a minuciosa análise sobre um conjunto de variáveis linguísticas e extralinguísticas de modo a detectar a sistematicidade da variação da concordância de número entre os constituintes do SN. Para tanto, focaliza os dados em duas perspectivas: a atomística (em que considera em separado cada elemento do SN) e a não atomística (em que considera o SN como unidade de análise). A autora mostra que o fenômeno estudado se apresenta como variação estável para um subgrupo de falantes e como mudança em progresso para outro, embora as variáveis mais relevantes atuem de maneira igual sobre as comunidades, o que constitui variação inerente.

Três variáveis interrelacionadas – posição linear do elemento no SN, classe gramatical e marcas precedentes – são discutidas em Scherre (1996), que, após verificar sua atuação isolada, propõe que estas sejam reduzidas a duas: (a) marcas precedentes em função da posição e (b) relação entre os elementos nucleares/não nucleares e posição dos elementos nucleares no SN. Com base na última destas novas variáveis, verifica que as classes não nucleares antepostas são mais marcadas que as pospostas, da mesma forma que elementos nucleares são mais marcados nas primeira e terceira posições do que na segunda.

Apoiando-se em resultados de sua tese e levando em conta a escrita padrão e a diversidade de contexto interacional, Scherre (1994b) apresenta evidências não só de que a variação na concordância nominal em PB é sistemática, mas também está bastante internalizada na mente do falante. Para tratar da influência do contexto interacional, apoia-se em Pereira (1993, apud Scherre, 1994b), que, analisando dados de um falante de nível superior de escolaridade, de classe média alta, administrador de uma fazenda, em três situações diferentes (interagindo com o patrão, com a família e com os empregados), mostra que a frequência da concordância diminui em cada uma das situações até chegar a quase zero, quando da interação com os empregados.

Ainda na perspectiva variacionista, Almeida; Brandão (1999) analisam os 321 SNs de três e quatro constituintes que ocorreram no *corpus* do trabalho de Almeida (1987), controlando os dados com base, entre outras, em uma variável em que se cruzaram a *distribuição do constituinte, sua posição no SN* e o *tipo de marcas precedentes*, variável essa que se mostrou a mais relevante para o cancelamento da marca de número. As autoras verificaram que os constituintes pré-nucleares tendem a ser enunciados com a marca de número e os constituintes nucleares e pós-nucleares a não apresentar a marca, por vezes de forma categórica e independentemente de haver ou não marcas formais e/ou semânticas anteriormente expressas. Levando-se em conta apenas a distribuição dos constituintes, isto é, a ordem em que ocorrem, deduziu-se que o princípio do paralelismo formal atua de uma maneira à esquerda do núcleo – marcas levam a marcas – e de outra, à sua direita – zeros

levam a zeros. Em síntese, antes do elemento nuclear, os constituintes tendem a receber o morfema de número. Depois do núcleo, em geral não-marcado, os constituintes tendem a não apresentar o morfema, sobretudo se próximos a ele, por influência do próprio núcleo.

Scherre (2005: 31-32) apresenta listagem bastante completa de textos inéditos e textos publicados que focalizam a concordância nominal, embora nos anos seguintes vários outros estudos tenham abordado o tema e ampliado os debates sobre a questão.

A partir de 2008, um grupo de pesquisadores brasileiros e portugueses, reunido no Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias*,² vem organizando *corpora* e realizando análises contrastivas no intuito de conhecer mais amplamente os padrões variantes de concordância nominal e verbal em português, de modo a contribuir não só para determinar a existência ou não de diferentes gramáticas em co-existência e em concorrência nos distintos espaços em que se fala o Português, mas também para discutir possíveis diferenças/semelhanças no processo histórico de constituição de suas variedades. É no âmbito desse projeto que se insere a análise comparativa que se segue.

2. Fatores atuantes no PB e no PST

Dos 1933 SNs que compõem o *corpus* do PB, e dos 1331 que constituem o do PST, 84%, respectivamente 1628 e 1120, têm o núcleo na segunda posição, a grande maioria formada por apenas dois constituintes, o que permite considerá-los como os SNs prototípicos da modalidade falada. Do quadro a seguir, que apresenta a posição linear dos núcleos no SN por variedade, ressalta o caráter homogêneo das duas amostras.

Posição do núcleo no SN	PB		PST	
	Número	Percentual no <i>corpus</i>	Número	Percentual no <i>corpus</i>
1 ^a	111	5,7 %	67	5,2 %
2 ^a	1628	84,2%	1120	84%
3 ^a	187	9,6 %	139	10,4%
4 ^a	7	0,5%	5	0,4%

Tabela 1 - Número de SNs por posição do núcleo por variedade

Os resultados gerais contrariaram a hipótese inicial: mesmo levando-se em conta que não foram considerados indivíduos com escolaridade inferior a cinco anos, esperava-se obter um maior índice de cancelamento, sobretudo em São Tomé, onde, segundo Hagemeijer (2009), a grande maioria da população é pelo menos bilíngue e onde coexistem línguas crioulas autóctones – o Santomé ou Forro e o Angolar – além do crioulo de Cabo Verde, do Português dos Tongas e de resquícios de línguas do grupo Banto.

² Cf. www.lettras.ufrj.br/concordancia

Variedade	Cancelamento da marca de número nominal	
	Ocorrências	Percentual
PB	338/3777	8,9%
PST	173/2612	6,6%

Tabela 2 – Índices gerais

Para melhor entender esse quadro, é importante atentar ao que se expõe na Tabela 3, em que se indicam as variáveis que se mostraram mais relevantes para a implementação do cancelamento.

PB	PST
Posição linear e relativa dos constituintes no SN	Nível de escolaridade
Nível de escolaridade	Posição linear e relativa dos constituintes no SN
Faixa etária	Animacidade do núcleo
Processo morfofonológico de formação de plural	Gênero
Animacidade do núcleo	Contexto fonológico subsequente
Significância: .001 Input: .039	Significância: .023 Input: .016

Tabela 3 - Variáveis atuantes para o cancelamento da marca de número por variedade

Em termos gerais, o *input* de cancelamento é baixo tanto na amostra do PB quanto na do PST, tendo em vista que foram considerados na análise todos os informantes, inclusive os que apresentam alto percentual de concordância canônica. No entanto, como se verificará no desenvolvimento da análise (e está patente nas variáveis selecionadas), esses índices encobrem não só uma diversificada realidade no que respeita aos grupos nelas representados, mas também os fortes condicionamentos estruturais que presidem à variação³. Observe-se que se alternam, em termos de importância, fatores sociais e estruturais, destacando-se, como os mais relevantes, respectivamente, *nível de escolaridade* e *posição linear e relativa dos constituintes no SN*.

A Tabela 4 mostra a mesma tendência nas duas amostras: os constituintes pré-nucleares apresentam menos cancelamento (na mesma proporção em ambas as variedades) do que os nucleares e os pós-nucleares, sendo a posição 1, no SN como um todo, a menos suscetível ao processo. Observam-se, em relação ao núcleo em segunda posição, pesos relativos (p.r.) idênticos (.74), diferenciando-se os elementos pós-nucleares dos nucleares pelo incremento dos índices de cancelamento.

³ Nas tabelas 4 e 7, apresentam-se exemplos referentes às duas variedades, o primeiro sempre relativo ao PB, o segundo, ao PST.

Posição linear e relativa dos constituintes no SN		PB			PST		
		Oco	%	P.R.	Oco	%	P.R.
PRÉ-NUCLEAR	1ª posição somos nós [o culpados] sei lá [essa influências]	19/1450	1,3	.17	7/1022	1	.15
	2ª/3ª posição com [as <i>minha</i> netinha] n[os <i>primeiro</i> momento]	6/155	3,9	.47	3/122	2	.50
NUCLEAR	1ª posição só se for [<i>sonho</i> bons] [<i>curso</i> noturnos de alfabetização]	4/111	3,6	.40	3/67	4	.57
	2ª posição [muitos <i>filho</i>] <i>que não entende</i> [as <i>dificuldade</i>] <i>que nós temos</i>	238/1628	14	.74	121/1120	10	.74
	3ª/4ª posição [todos os <i>político</i>], não [todas essas <i>língua</i>]	18/194	5,1	.68	14/144	9,7	.76
PÓS-NUCLEAR	2ª posição falar [palavras <i>errada</i>] [trabalhos <i>específico</i>]	9/95	9,5	.72	10/58	17,2	.91
	3ª posição tem [esses transportes <i>alternativo</i>] passar [os tempos <i>livre</i>] em casa	34/109	31,2	.90	10/60	16,7	.88
	4ª/5ª posição [os filhos tudo <i>formado</i>] [as pessoas mais <i>velha</i>]	10/35	28,6	.89	5/19	26,3	.94

Tabela 4 - Atuação da variável *Posição linear e relativa no SN* para o cancelamento da marca de número, por variedade.

Naro; Scherre (2003b), em estudo de tempo real de curta duração com amostras da década de 1980 e de 2000 referentes a indivíduos urbanos cuja escolaridade varia de 4 a 11 anos, afirmam que, apesar de, entre uma fase e outra, terem aumentado percentualmente os índices de concordância nominal (de 68% para 81%, no geral), o padrão estrutural que rege a variação mantém-se o mesmo: por exemplo, elementos à esquerda na 1ª. posição continuam a

favorecer a marca explícita (p.r. .92, em 80 e .85 em 2000), núcleos na segunda posição continuam a desfavorecê-la (p.r. .16, em 80, e .21, em 2000) e assim por diante.

Pelo que os dados do presente estudo demonstram, os padrões registrados no PB repetem-se no PST, como fica claro no gráfico 1, que diz respeito ao cancelamento da marca por atuação da posição relativa dos constituintes⁴. Note-se que, embora o *input* do cancelamento seja menor na fala de São Tomé, talvez, por isso, a pressão exercida pelos fatores seja bem mais significativa do que na do Brasil.

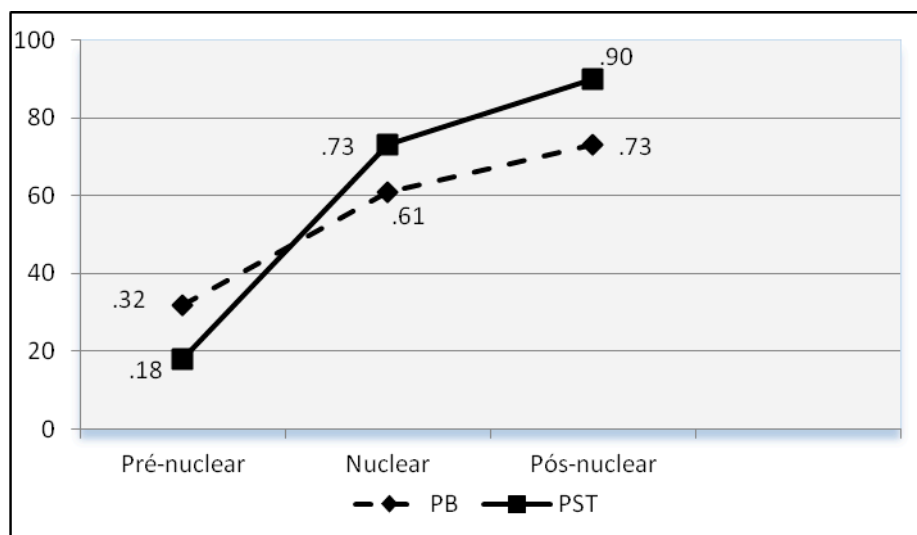


Gráfico 1 - Cancelamento da marca de número no SN segundo a posição relativa do constituinte (pesos relativos).

Outra variável significativa em ambas as amostras foi o nível de escolaridade, que, no PST, se mostrou a mais relevante. Nessa variedade, a probabilidade de cancelamento da marca de número vai gradativamente diminuindo em função da maior exposição do indivíduo à educação formal: a diferença de peso relativo entre os menos escolarizados e os mais escolarizados chega a .67. Já no PB, o desempenho de indivíduos com nível superior (p.r. .24, também idêntico ao que se obteve em PST), contrasta com o dos indivíduos de nível fundamental e médio, sendo estes últimos os que mais se destacam quanto ao cancelamento.

Nível de escolaridade	PB			PST		
	Oco	%	P.R.	Oco	%	P.R.
5 a 8 anos (fundamental)	101/1175	8,6	.62	125/531	23,5	.91
9 a 11 anos (médio)	193/1013	19,1	.76	34/942	3,6	.51
12 a 15 anos (superior)	44/1589	2,8	.24	14/1139	1,2	.24

Tabela 5 – Atuação da variável *nível de escolaridade* para o cancelamento da marca de número no SN, por variedade

⁴ Os pesos relativos foram obtidos em rodadas que tiveram significância .001, para o PB, e .018 para o PST.

No PB, em que também a faixa etária é relevante (cf. Tabela 6), são os indivíduos mais velhos os mais predispostos ao cancelamento (p.r. .64), enquanto os mais jovens tendem a adotar a marca de concordância, com índices bastante aproximados (18-35 anos: p.r. .48; 36-55 anos: p.r. .41).

Faixa etária	Oco	%	P.R.
18-35 anos	96/1122	8,6	.48
36-55 anos	104/1580	6,6	.41
56-75 anos	138/1075	12,8	.64

Tabela 6 - Atuação da variável *faixa etária* para o cancelamento da marca de número no SN, no PB

Para melhor aferir a variável *nível de escolaridade*, procedeu-se ao seu cruzamento com faixa etária, cujos resultados, apresentados nos Gráficos 2 e 3, permitem as seguintes generalizações referentes a ambas as variedades:

- nas diferentes faixas etárias, os indivíduos de nível superior de instrução apresentam os menores índices de cancelamento, o que sugere que, nesse grupo, a regra de concordância nominal tem caráter semicategórico, nos termos definidos por Labov (2003);
- os indivíduos mais velhos com menor grau de escolaridade (fundamental e médio) detêm os maiores índices de cancelamento, embora em São Tomé também se destaquem nesse sentido os da faixa etária intermediária de nível fundamental;
- as curvas referentes aos indivíduos de nível fundamental e médio sugerem um movimento no sentido de incorporação da norma padrão de concordância.

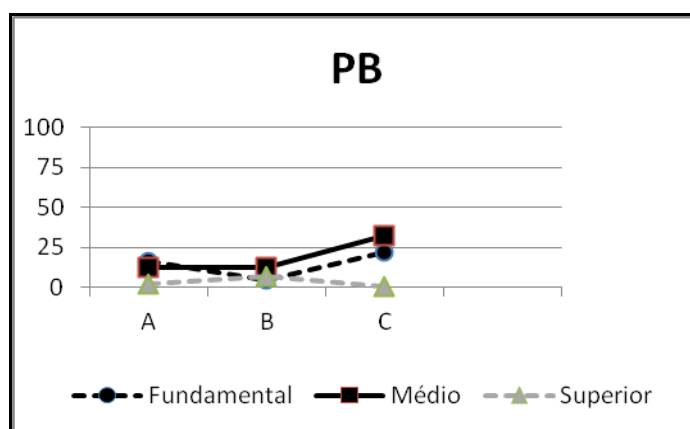


Gráfico 2 - Cruzamento de nível de escolaridade e faixa etária: percentuais de cancelamento

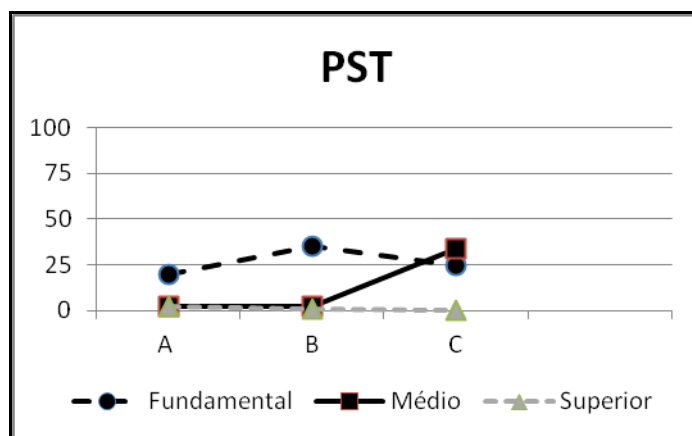


Gráfico 3 - Cruzamento de nível de escolaridade e faixa etária: percentuais de cancelamento

Quanto à variável *animacidade*, sistematicamente selecionada nas análises do PB e do PST, se poderia aplicar a afirmação de Scherre (1988) de que, embora os resultados a ela pertinentes sejam “consistentes” não apresentam “diferenças probabilísticas fortes”.

Aqui optou-se por considerar em conjunto núcleos com os traços [+/- humano] e [+ animado] contra núcleos com traços [- humano] e [- animado], em função do pequeno número de dados com os traços [- humano] e [+ animado], quatro e nove, respectivamente, nas amostras do PB e do PST.

De qualquer forma, a hipótese confirma-se: vocábulos com traço [+ animado], em que predomina o traço [+ humano], são menos sujeitos ao cancelamento do que os de traço [- animado], embora com índices probabilísticos próximos à neutralidade.

Animacidade do núcleo	PB			PST		
	Oco	%	P.R.	Oco	%	P.R.
[+ animado]						
[as <i>menina</i>] tiraram retrato	95/838	11,4	.43	36/489	7	.40
manter [os <i>filho</i>] na escola						
[- animado]						
na minha casa tem [três <i>geladeira</i>]	166/1095	15,2	.55	102/842	12	.55
comprar [uns <i>medicamento</i>]						

Tabela 7- Atuação da variável *animacidade do núcleo* para o cancelamento da marca de número no SN, nas duas variedades

3. Fatores atuantes só no PB

Além dos fatores relativos à faixa etária, em PB, mostrou-se atuante a variável *processo morfofonológico de formação de plural* que busca dar conta da atuação do princípio de saliência fônica (LEMLE; NARO, 1977), segundo o qual formas menos marcadas, neste caso aquelas em que a diferença entre a forma singular e plural reside apenas na presença do morfema de número – como em *filh[ɔ]/filh[ɔj]* – são mais predispostas à perda da marca do que formas como *[o]vo/[ɔ]vos*.

Os itens menos marcados do ponto de vista fônico apresentam maior tendência ao cancelamento da marca, o que se tem comprovado em diversas outras pesquisas.

Processo morfofonológico de formação do plural		Número	%	P.R.
Menor saliência fônica	<i>Itens terminados em vogal oral ou nasal</i>	313/3403	9,2	.52
	filho-filhos/homem/homens			
Maior saliência fônica	<i>Itens terminados em -R, S, L e -ÃO e com marca dupla</i>	25/374	6,7	.30
	cantor/cantores			
	mês-meses/Atual-atuais			
	cartão-cartões/[o]vo-[ɔ]vos			

Tabela 8 - Atuação da variável *Processo morfofonológico de formação do plural* para o cancelamento da marca de número no SN, no PB

4. Fatores atuantes só no PST

No PST, gênero e contexto fonológico subsequente também se mostraram salientes para o cancelamento como se observa nas Tabelas 9 e 10.

Gênero	Oco	%	P.R.
Masculino	18/1607	7	.54
Feminino	55/1005	5	.43

Tabela 9 – Atuação no PST

Contexto fonológico subsequente	Oco	%	P.R.
Vogal/consoante	95/1984	4	.47
Pausa	78/628	12	.58

Tabela 10 – Atuação no PST

5. Reflexões sobre os resultados obtidos

As convergências, neste breve estudo, são evidentes, a começar pelo próprio formato das amostras que serviram de base às análises. Embora os dois *corpora* apresentem uma diferença de 602 SNs, seu perfil, praticamente idêntico (cf. Tabela 1), deu ensejo a que se pudesse realizar a análise segundo os mesmos parâmetros. Os dois *corpora* têm, ainda, em comum o fato de serem representativos da fala de indivíduos que, embora aqui distinguidos fundamentalmente pelo nível de escolaridade, vivem em áreas urbanas, em constante interação interdialeto e interlinguística, como no caso de São Tomé, com grupos sociais os mais diversos. Tudo isso permitiu que se chegasse a resultados que suscitam antes a formulação de hipóteses do que respostas conclusivas.

O português, em São Tomé, língua oficial desde 1975, coexiste com outras línguas, entre as quais o Forro, crioulo de base portuguesa que, segundo Afonso (2008, p.74), apesar de falada por 73% da população, é fortemente estigmatizada, o que tem sua origem na política adotada pelo colonizador ao longo do tempo e hoje introjetada na mente e na prática de seus próprios utentes, sobretudo da elite, que parece desconsiderar o caráter multilíngue da sociedade são-tomense. Alguns dos informantes da pesquisa mencionaram a obrigatoriedade de falar o português, mesmo nas situações familiares de intercomunicação, por imposição dos pais, temerosos de que os filhos não se expressassem de acordo com as normas de referência, as do português europeu, e assim não pudessem auferir as poucas oportunidades de ascensão socioeconômica num país em construção.

A realidade linguística do Brasil vem sendo caracterizada como variável, heterogênea e polarizada (LUCCHESI, 1998, p. 74), o que se depreende de diversos trabalhos, como o de Almeida (1997), que, focalizando a concordância nominal, na fala de indivíduos analfabetos ou de baixo índice de escolaridade (1 a 4 anos) em áreas rurais do Estado do Rio de Janeiro, indica que, dos 2.865 SNs selecionados da fala dos 78 informantes que contribuíram para a sua amostra, 2032 são compostos apenas por dois constituintes e, ainda, que o cancelamento da marca de número no núcleo é altamente produtivo: 87% e 86%, respectivamente, na segunda e na terceira posição. Já Campos; Rodrigues (1993), com base na fala registrada em cinco grandes cidades brasileiras⁵, indicou que, no Rio de Janeiro, num total de 735 vocábulos suscetíveis de flexão, apenas 2%, apresentavam cancelamento. Estudos recentes em tempo real de curta duração em áreas urbanas (NARO; SCHERRE, 2003), como já se afirmou, têm chamado a atenção para o incremento das taxas globais de concordância, não obstante os condicionamentos que presidem à variação permanecerem nos mesmos patamares. Já em artigo de 1994b, Scherre observava que

o fenômeno da variação de número no português do Brasil pode ser caracterizado como um caso de variação linguística inerente, tendo em vista que ocorre em contextos linguísticos e sociais semelhantes e apresenta tendências sistemáticas de variação altamente previsíveis (p. 37).

Baxter (2009), focalizando os dialetos de duas comunidades afro-descendentes, uma brasileira, outra são-tomense, depois de apontar uma série de “fatos linguísticos e sociolinguísticos compartilhados” pelas duas variedades (p. 270-271), postula que,

no seu desenvolvimento, o sistema começa com um PL singleton, mas depois desenvolve características de concordância parcial nos elementos mais próximos do fulcro: elementos do DP antes do NP (refletindo um crescimento estrutural nesse componente) e para a direita (no NP propriamente dito) (p. 292)

Convergências entre as duas variedades parecem também ocorrer no que respeita à concordância verbal. Estudo, ora em desenvolvimento, sobre essa variável, com base na fala dos mesmos informantes, vem demonstrando que se repetem os baixos índices gerais de cancelamento nas duas variedades.

⁵ Os índices de cancelamento da marca de número nominal na variedade culta foram de 2% no Rio de Janeiro, 7% em Porto Alegre, 6% em Salvador, 1%. Em Recife, não houve ocorrências de apagamento.

A questão, no entanto, é mais complexa, pois, na modalidade oral, a concordância nominal, no PB e, ao que tudo indica, no PST, situa-se num *continuum* que opõe, em comunidades urbanas, de um lado, os indivíduos [- letrados] e de outro, os [+ letrados], isso sem levar em consideração outros contínuos, como os *de urbanização* e *de monitoração estilística*, conforme a proposta de Bortoni-Ricardo (2004). Há de se considerar, ainda, a sócio-história de cada uma das variedades, divergente em vários aspectos.

Se já muito se conhece sobre a dinâmica do PB, o mesmo não ocorre com a do PST. Seria importante, por isso, empreender outras pesquisas sobre esta e outras variáveis não só no âmbito do PST, mas também no do Português de Moçambique, de Angola, de Cabo Verde, de modo a apreender convergências/divergências entre as variedades, verificando em que medida mais se aproximam ou se distanciam dos padrões do Português Europeu e mais ou menos divergem/convergem entre si. Acredita-se que, assim, seria possível ter mais clareza quanto ao processo que estaria na base da diferenciação entre o PB (ou entre as demais variedades) e o PE, se a transmissão linguística irregular ou a deriva românica.

Conclusão

Neste estudo, focalizou-se a concordância no âmbito do SN em variedades urbanas do PB e do PST, no intuito de verificar se haveria convergências e/ou divergências no que tange aos fatores que motivam a não ocorrência da marca de número plural.

Conclui-se do que se expôs ao longo da análise, que o PB e o PST, no que respeita à concordância nominal, compartilham fatos linguísticos e sociais em suas variedades urbanas, havendo em comum:

- (a) padrões variáveis que vão da categoricidade da marcação de plural em todos os constituintes à marcação apenas no primeiro deles, não se observando estrutura(s) que seja(m) privativa(s) de uma das variedades;
- (b) fortes restrições não só de ordem estrutural, como a que diz respeito ao cancelamento da marca de plural segundo a posição linear e relativa do constituinte no SN, mas também de natureza social, entre as quais se destaca a maior ou menor exposição do indivíduo à educação formal;
- (c) índices bastante aproximados no que concerne à atuação dos fatores que compõem três das cinco variáveis que se mostraram atuantes para o cancelamento;
- (d) a significativa atuação dos fatores sociais: *nível de escolaridade* em ambos os casos, *faixa etária* no PB e *gênero* no PST;
- (e) o caráter semicategórico da aplicação da marca de plural entre os indivíduos mais escolarizados, sobretudo em São Tomé, em que a norma de referência é o Português Europeu;
- (f) indícios de que a marca de número tem como *locus* preferencial a margem esquerda do SN, onde se concentram os determinantes, e que sua maior ou menor incidência nos demais constituintes depende do maior ou menor grau de escolaridade do falante ou, em outros termos, de seu maior ou menor conhecimento das normas de prestígio.

Embora, como se indicou no item 1, haja, em relação ao PB, um número bastante significativo de trabalhos sobre o tema aqui focalizado, em diferentes perspectivas teóricas, pouco ainda se sabe, do ponto de vista mais estritamente sociolinguístico, sobre as variedades africanas do Português. Nesse sentido, a análise realizada, além de confirmar os resultados obtidos em outros estudos sobre o PB, não só contribui para o conhecimento de uma dessas

variedades, mas também fornece outros elementos para discutir o processo que estaria na base da diferenciação entre o PB e o PE.

ABSTRACT: The article focuses on the number agreement in noun phrase in urban varieties of Portuguese spoken in Brazil and in Sao Tome, in order to determine the factors that contribute to the coexistence of variant patterns found in them. The results, coming from analysis based on the assumptions of Variationist Sociolinguistics, show that there are significant similarities between the two varieties, among them the strong structural and social constraints that guide the deletion of the number morpheme.

Keywords: number agreement; noun phrase; Portuguese of Brazil; Portuguese of Sao Tome; urban variety.

Referências

ALMEIDA, E. M. *A variação da concordância nominal num dialeto rural*. 1997. 145f. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 1997.

AMARAL, A. *O dialeto caipira*. 3 ed. São Paulo: Hucitec; Secretaria de Cultura, Ciência e Tecnologia, 1976 [1920]. 197p.

BAXTER, A. A concordância de número. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (Org.) *O português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 269-293.

BRANDÃO, S. F.; ALMEIDA, E. M. Ainda sobre a concordância no âmbito do sintagma nominal. In: ESTUDOS DA LINGUAGEM: RENOVAÇÃO E SÍNTESE. VII CONGRESSO DA ASSEL-RIO. *Anais...* Rio de Janeiro, 1998. Rio de Janeiro: Associação de Estudos da Linguagem do Rio de Janeiro (ASSEL-Rio), 1999. p. 835-843.

BORTONI-RICARDO, S. M. *Educação em língua materna: a sociolinguística na sala de aula*. São Paulo: Parábola Editorial, 2004. 112p.

BRAGA, M. L. *A concordância de número no sintagma nominal no Triângulo Mineiro*. 1977. 88f. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica, Rio de Janeiro, 1977.

_____; SCHERRE, M. M. P. A concordância de número no SN na área urbana do Rio de Janeiro. In: ENCONTRO NACIONAL DE LINGÜÍSTICA, 1, 1976, Rio de Janeiro. *Anais ...* Rio de Janeiro: PUC, 1976. p. 464-477.

CAMPOS, O. G. L. de S.; RODRIGUES, A. C. S. Flexão nominal: indicação de pluralidade no sintagma nominal. In: ILARI, R. (Org.) *Gramática do português falado*. 2. ed. Campinas-SP: Editora da UNICAMP, 1993. vol. 2: Níveis de análise linguística. p. 111-134.

GUY, G. R. *Linguistic variation in brazilian portuguese: aspects of the phonology, syntax, and language history*. 1981. 391f. Ph. D. Dissertation - University of Pennsylvania, Philadelphia, 1981.

HAGEMEIJER, T. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, v.1, n. 1, p.1-27, 2009.

LABOV, W. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Ed.) *Sociolinguistics: the essential readings*. Blackwell Publishing, 2003. p. 234-250

LEMLE, M.; NARO, A. J. *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, D. O conceito de transmissão linguística irregular e o processo de formação do Português do Brasil. In: RONCARATI, C.; ABRAÇADO, J. (org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003. p. 272-284.

_____. A constituição histórica do português brasileiro como um processo bipolarizador: tendências atuais de mudança nas normas culta e popular. In: GROBE, S.; ZIMMERMANN, K. (Eds.), *'Substandard' e mudança no português do Brasil*. Frankfurt: TFM, 1998. p. 73-99.

NARO, A. J.; SCHERRE, M. M. P. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras, 2003a. p. 285-302.

_____. Estabilidade e mudança linguística em tempo real: a concordância de número. In: PAIVA, M. C. de; DUARTE, M. E. L. (Org.). *Mudança linguística em tempo real*. Rio de Janeiro: FAPERJ; Contra Capa, 2003b. p. 47-62.

PONTE, V. M. L. 1979. 215f. *A concordância nominal de uma comunidade de Porto Alegre*. Dissertação (Mestrado em Linguística) – Pontifícia Universidade Católica, Porto Alegre, 1979.

RODRIGUES, A. N. *O dialeto caipira na região de Piracicaba*. São Paulo: Ática, 1974. 222p.

SCHERRE, M. M. P. 1978. *A regra de concordância de número no sintagma nominal em português*. 158f. 1978. Dissertação (Mestrado em Letras) - Rio de Janeiro, Departamento de Letras e Artes, Pontifícia Universidade Católica, 1978.

_____. *Reanálise da concordância de número em português*. 1988. 2 v. Tese (Doutorado em Linguística). – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 1988.

_____. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa*, Lisboa, n. 12, p. 37-49, 1994a.

_____. Grau e formalidade léxica na concordância nominal em português In: SEMINÁRIO DE PÓS-GRADUAÇÃO EM LETRAS - PESQUISA, 3, 1988, Rio de Janeiro. *Anais...* Rio de Janeiro, UFRJ/FL, 1994b. p. 147-166.

_____. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: SILVA, G. M. de O.; SCHERRE, M. M. P. (Orgs.). *Padrões sociolinguísticos-análises de fenômenos variáveis do português falado no Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, 1996. p. 85-117.

_____. *Doa-se lindos filhotes de poodle: variação linguística, mídia e preconceito*. São Paulo: Parábola Editorial, 2005. 160p.

VEADO, R. M. A. *Comportamento linguístico do dialeto rural*. Belo Horizonte: UFMG/PROED, 1982. 109p.

RECEBIDO EM 29/10/2010 - APROVADO EM 13/05/2011