

SOBRE AS ORIGENS DO PORTUGUÊS POPULAR DO BRASIL

Anthony Julius NARO (Universidade Federal do Rio de Janeiro)

Maria Marta Pereira SCHERRE (Universidade Federal do Rio de Janeiro - Universidade de Brasília)

ABSTRACT: In our view, at the time the Portuguese language was transferred from Portugal to Brazil it carried along a centuries old drift, partially inherited from Latin and earlier languages, that, in its new context, interacted with a series of factors that may have reinforced aspects of the original drift. From the outset, one of the strongest factors in Brazil was pidginization, in the form of the Tupi based "lingua geral" as well as the acquired code known as "lingua de preto" in Portugal and used with Indians in the New World. Both of these pre-date the Portuguese presence in Brazil, probably going back several centuries. Later, African based pidgins may have come into play due to the predominance of people from certain regions of Africa in the north and south of Brazil, although there is no evidence to point to the existence of a Portuguese based pidgin or creole peculiar to descendants of Africans brought to Brazil under slavery. In any case, this particular pidgin's effect on Portuguese would not be distinguishable from that of the other non-Portuguese based pidgins. Throughout the history of the Portuguese language in Brazil, contact with populations of diverse linguistic origins has caused the language to develop differently than in Portugal, but the ultimate source of the changes that occurred in Brazil can be traced back to Portugal.

0. Introdução

Neste artigo, temos como objetivo principal fazer algumas reflexões a respeito da origem do português popular do Brasil, apresentando evidências históricas, linguísticas e sociais. A nossa posição se resume na frase "confluência de motivos": vemos a atração de forças de diversas origens - algumas oriundas da Europa, outras da América, outras, ainda, da África - que juntas se reforçaram para produzir o português popular do Brasil. Hoje em dia esta posição parece-nos quase consensual (Silva 1992, p.86, Guy 1989, p.16; Bortoni-Ricardo 1981, p.81).¹ Não atribuímos um papel exclusivo a um

suposto pidgin ou crioulo de base lexical portuguesa, entendido como uma língua falada pela comunidade negra e estruturalmente diferente da de outras etnias.² Parece-nos oportuno lembrar também o papel dos índios, o das demais etnias presentes no Brasil e até a contribuição pidginizante dos primeiros colonos portugueses, além de outras forças em interação com a deriva secular trazida da Europa.

1. Argumentação histórica

Para a história da língua portuguesa no Brasil, existe uma documentação razoavelmente rica e variada que vai desde relatos de viajantes, missionários e autoridades governamentais, a partir do século XVI, a representações de fala de alguns grupos populacionais em peças, jornais etc., a partir, principalmente, do século XIX. O mais completo levantamento deste material ainda é o de Serafim da Silva Neto, em *Introdução ao Estudo da Língua Portuguesa no Brasil*, cuja primeira edição data de 1950 e de quem tomamos emprestadas quase todas as citações a seguir. Estas fontes revelam com muita segurança o predomínio quase total da "língua geral" - um pidgin ou koiné simplificado de origem tupi - até pelo menos o início ou meados do século XVIII. Todos os grupos étnicos - índios de diversas tribos, europeus, e africanos - sabiam se comunicar usando este sistema tupi, sem dúvida com conhecimentos mais ou menos profundos segundo as circunstâncias individuais de cada um. A partir do século XVIII, entretanto, a língua portuguesa começa a se espalhar entre a população brasileira até chegar à situação atual de seu predomínio maciço mesmo entre populações com pouca ou nenhuma ascendência lusa (cf., também, Rodrigues, 1983).

Apesar da riqueza das fontes disponíveis a respeito da fala dos índios e dos brancos, é escassa qualquer evidência documental específica quanto ao português ou outras línguas faladas pelos africanos no Brasil, além de algumas menções breves de africanos que não dominavam o português, mas apenas a língua geral tupi ou línguas africanas. A documentação não transmite nem a mais leve suspeita de que a língua portuguesa falada pelos brasileiros descendentes dos primeiros cativos africanos fosse diferente da fala de brasileiros de outras origens étnicas (mantidas, naturalmente, as semelhanças socio-econômicas relevantes). Nesse aspecto, a situação de então parece espelhar a de hoje: nos dois casos os observadores - nativos ou estrangeiros - não conseguem reconhecer qualquer traço lingüístico

associado exclusivamente com a etnia afro-brasileira. Pode existir um "Black English" nos EUA; por aqui se existiu um "Português Negro" suas diferenças em relação ao português geral eram tão sutis que fugiam aos ouvidos dos observadores.

É importante notar que não estamos querendo afirmar que os falantes nativos de línguas africanas ou indígenas do Brasil tenham adquirido a língua portuguesa com perfeição nativa, nem que os falantes nativos de português não tenham introduzido em sua fala traços africanos, indígenas ou pidginizantes à maneira da língua geral ou até da chamada "língua de preto", tão arraigada em Portugal pelo menos desde o século XV e usada para comunicação com africanos tanto na África como na Europa. Pelo contrário, para o século XVII, temos o depoimento de Pe. Antônio Vieira, que nos diz que as várias nações da Ásia sabem falar a língua portuguesa, "mas cada uma a seu modo, como no Brasil os de Angola e da terra". Um século antes, João de Barro tinha dito: "podemos dizer que as nações de "África, Guiné, Ásia, Brasil barbarizam quando querem imitar a nossa [língua]", onde por "barbarizam" ele entende "não podem formar a linguagem". E no século XVIII um viajante diz dos índios caiapó que, embora soubessem falar corretamente o português, o faziam com "um sotaque tão fastidioso e ingrato, que fazia fugir". Mas nem por isso os brasileiros não indígenas deixavam de ser influenciados pelos idiomas da terra, o que não surpreende dada a expansão da língua geral no país.

O quadro lingüístico inicial que surge então é o de uma comunidade em que as línguas dos diversos grupos se influenciavam, principalmente através do aprendizado imperfeito de falantes não-nativos adultos. Sabemos que o predomínio inicial - a língua geral - era pidginizante. Do português deste período temos apenas um texto publicado em 1620, quando a população de origem africana no Brasil ainda era muito pequena, em que um missionário imita a fala dos índios. Este "português de índio" do Brasil é, entretanto, muito semelhante, senão idêntico, ao "português de preto" da Europa e da África, mostrando traços como aplicação variável de concordância nominal e verbal, pronomes plenamente acentuados ("mim", "bozo"), acréscimo de vogais finais, redução de palatais sonoras a yod, troca entre *l*, *r*, *d*, etc. Uma influência direta dos africanos sobre os índios teria sido impossível no início do século XVII, devido ao predomínio da língua geral e ao pequeno número de africanos nessa época. Parece bem mais provável que este texto seja reflexo direto de uma transferência do pidgin da

Europa para a América, operada pelos próprios portugueses. Completa-se assim o quadro lingüístico dos primeiros séculos do Brasil: predomínio do pidgin tupi, influência mútua das diversas línguas no contexto do aprendizado imperfeito e elementos pidginizantes vindos da Europa.

A este quadro se juntou, em números sempre maiores, a população de origem africana. A única informação segura que temos a respeito da situação lingüística dos primeiros africanos chegados ao Brasil é a de que falavam línguas africanas, e, em alguns casos e em diversos graus, adquiriram conhecimentos da língua geral ou do português. Falta qualquer informação sobre o tipo - pidginizado ou não - destas línguas. Entretanto, dado o fato de que as populações africanas eram muito concentradas geograficamente, com falantes de línguas Bantu para o sul e os da região do Yorubá mais para o norte, parece inevitável a existência, entre os africanos no Brasil, de pelo menos duas línguas gerais de base² africana, de forma semelhante ao sistema de base tupi atestado. De fato, esta é a hipótese clássica na lingüística brasileira, sugerida tão cedo quanto o início deste século. Nelson Rodrigues (1932), em seu *Os africanos no Brasil*, avança a hipótese de dois sistemas de base africana, resultantes do contato de várias línguas africanas entre si e dessas com o português: um pidgin de base Yorubá mais ao Norte e outro de base Kimbundu mais ao Sul (cf, também, Rodrigues, 1983, p.30-3).

Com a investigação recente na comunidade do Cafundó, na área caipira no Estado de São Paulo, agora temos evidências em favor das idéias de Nelson Rodrigues. Esta pequena comunidade, consistindo de cerca de 80 pessoas, tem suas origens diretas na doação de um pedaço de terra a dois ex-escravos, ancestrais da população atual, pelos seus ex-proprietários, um pouco antes da Abolição em 1888. Então, as origens da comunidade do Cafundó são precisamente na época da maior influência africana no Brasil, quando a imigração européia em ampla escala estava apenas começando. Embora esta comunidade seja separada fisicamente da população da redondeza, o pessoal do Cafundó participa da economia local de várias formas e mantém contatos sociais com pessoas de várias raças, cores e status sociais. Sua língua nativa parece ser o português caipira, indistinguível do dos seus vizinhos, mas eles também falam uma língua (africana), usada entre eles mesmos como um código secreto. Em termos grosseiros, esta língua africana pode ser considerada como o português caipira típico, com quase todos os

verbos e nomes substituídos por formas de base Bantu. Em contraste com os pidgins melhor conhecidos, estereotipados como gramática "Africana" com léxico português, o sistema Cafundó seria gramática portuguesa com léxico africano (Bantu). A origem desta situação *sui generis* pareceria ser a africanização superficial do português popular do Brasil local do período, através da relexificação para os propósitos de comunicação intragrupal, a partir da língua geral Bantu acima referida.

Projetando para o passado o que ocorre com o Cafundó, é fácil entender porque a literatura não contém alusão a um pidgin de base portuguesa - provavelmente não havia; parece mais verossímil que os africanos falavam variantes do português popular do Brasil local do período ou um pidgin de base africana que, com o correr do tempo, caiu em desuso exceto para os casos de comunicação secreta. Esta abordagem tem a vantagem de explicar a pesada influência lexical das línguas africanas no português popular do Brasil

Parece então improvável que tenha existido no Brasil uma língua pidgin ou crioula de base lexical portuguesa associada predominantemente com a etnia afro-brasileira ou ameríndia. Tal língua era dispensável, dada a existência de outras "línguas gerais", de bases não européias, que já preenchiam as necessidades comunicativas da população. Mas não podemos deixar de suspeitar que a pidginização em si, quase endêmica no Brasil desde o início da Colônia (e antes, tanto no caso da "língua de preto" da Europa e, provavelmente, das línguas gerais tupi e africanas), tenha influenciado no desenvolvimento do português brasileiro.

2. Motivações históricas européias

Um dos fenômenos mais frequentemente citados como sendo de origem crioula no português do Brasil é a variabilidade dos sistemas de concordância nominal e verbal (Guy, 1989). As línguas crioulas típicas, bem como certas línguas da África ocidental, costumam apresentar uma única forma lexical que não admite qualquer modificação para indicação de noções subsidiárias do tipo pessoa, gênero, número, tempo, aspecto etc. Segundo este ponto de vista, a perda (parcial ou total) dos mecanismos de concordância no Brasil seria o resultado da "africanização" do português.

É perfeitamente possível, entretanto, que a língua portuguesa já possuísse o embrião do novo sistema mais analítico antes de sair da Europa. Tal estado de coisas se torna bastante plausível dada a deriva secular das línguas românicas, e indo-européias de maneira geral, em direção à uniformização morfológica, com a sobrevivência apenas das formas 'irregulares' mais salientes.

A variação na concordância verbal tem um componente que parece puramente fonológico: quando o plural *comem* [ko'm ẽ] se reduz ao singular *come* ['kom:t] a única diferença existente pode ser a perda da nasalização da vogal não acentuada final. Tal regra existe atualmente no português do Brasil e opera variavelmente também sobre formas não verbais do tipo *homem*, *ontem* etc. Por outro lado, quando *comeram* [ko'merũ] é substituído por *comeu* [ko'meu] a diferença consiste na substituição de uma desinência *-eram* por outra *-eu* e o fenômeno não mais parece fonológico. Existe evidência empírica que sugere que em etapas anteriores do desenvolvimento da concordância verbal o tipo *comem/come*, em que atua a regra fonológica da desnasalização, liderava maciçamente a redução da concordância (Naro, 1981, p.93). Dai conclui-se que a redução morfológica da concordância é um desenvolvimento mais tardio, criado a partir da generalização da redução fonológica. De fato, esta primeira etapa - a desnasalização - existe na fala popular de Portugal, principalmente na região de Entre-Douro-e-Minho, onde, segundo Leite de Vasconcellos (1972, 86-7), a regra atua tanto em nomes (*virgem* como *birge*) como em verbos (*vertem* como *bérte*). Temos aí uma origem europeia da redução da concordância verbal, sem qualquer influência exterior necessária, africana ou ameríndia. Esta regra é, aliás, bem antiga: a omissão da nasal final é freqüente nos textos medievais portugueses e até em latim clássico a nasal final era fraca (Grandgent, 1962, p.193). Omitia-se a nasal tanto em inscrições pré-clássicas como em inscrições plebéias tardias (Sturtevant, 1940, p.151), tendo sido restaurada apenas durante a fase da língua clássica escrita.

Como vimos, há evidências de que o fenômeno da desnasalização não é específico do português do Brasil. Admitindo que a mudança lingüística que envolve a concordância verbo-sujeito tenha se iniciado na fonologia, precisamente através da desnasalização, concluímos que as suas origens remontam pelo menos até aos tempos pré-clássicos. Resta saber se podemos aplicar o raciocínio acima à concordância nominal. Embora sejam raras as menções à ausência do *-s* final no

português Europeu, temos evidências históricas do comportamento variável do -s, desde o latim antigo até as línguas românicas ocidentais modernas.

Para o português Europeu, vamos encontrar, nas palavras de Leite de Vasconcellos, evidências de enfraquecimento do -s de forma geral e queda do -s final em circunstâncias bem especiais. Segundo este autor, pronuncia-se o -s enfraquecido diante de consoante surda e no final de palavras no Centro e no Norte de Portugal (cf. 1970, p.97). Observa também que, em Estremadura, ocorre frequentemente a substituição do -s por -r, em sílabas átonas diante de uma consoante sonora. Além disso, afirma que "na pronúncia corrente -s (-z) se perde diante de -r (=rr)" e se neutraliza completamente diante de fricativas não labiais (cf. 1970, p.98).

Na história das línguas românicas em geral, temos muitas evidências de queda ou enfraquecimento do -s final. Já nos dialetos itálicos pré-românicos encontramos freqüente omissão do -s final e, nas inscrições latinas primitivas, encontra-se o -s omitido livremente. Na época clássica, porém, e continuando até o século II o -s reaparece, para então desaparecer de novo (cf. Grandgent, p.190). Mesmo no latim clássico o -s era apagado sob certas circunstâncias, embora Cícero taxasse esta pronúncia de "subrusticum" (cf. Sturtevant, p.161). Desses fatos conclui-se que o -s final dificilmente era pronunciado na língua popular de qualquer fase da história do Latim. Grandgent (cf. p.91), entretanto, é da opinião que o -s deve ter sido reforçado na Gália, Espanha e outras regiões devido a "primitivos hábitos lingüísticos dos indígenas" e assim se salvou em certas regiões. Como quer que seja, mesmo na Romania Ocidental constataram-se, desde o início, vacilações na realização de -s, que cai sob determinadas circunstâncias em sardo e sobretudo em francês, onde se criou toda uma nova sistemática de marcação de plurais. No francês, como se sabe, a indicação explícita do plural se faz através dos artigos (*les tables*), dos possessivos (*mes tables*), dos números (*deux tables*), dos demonstrativos (*ces tables*) e ainda de outros elementos gramaticais que se concentram na parte inicial do sintagma nominal. Estas mesmas características de localização da marca formal do plural no início do sintagma nominal também se encontram em diversos pidgins/crioulos, bem como em algumas línguas africanas (cf. Guy, 1989, p.233). Adiante discutimos a questão da localização da marca de plural no sintagma nominal no português popular do Brasil.

Os fatos acima arrolados mostram que o processo da queda do -s final no português do Brasil pode ter tido seu início no português dialetal da Europa, que, por sua vez, estava apenas dando continuidade a uma deriva pré-românica. Conseqüentemente, não é implausível supor que o impulso inicial do processo de perda da concordância nominal se situe também em fenômenos fonológicos trazidos da Europa, à semelhança da concordância verbo/sujeito.

3. Sobre o papel da posição linear na concordância nominal

A posição linear que o elemento ocupa no sintagma nominal tem sido considerada a variável lingüística mais importante nos diversos trabalhos variacionistas que analisaram a concordância nominal de número no português falado no Brasil (cf. Braga & Scherre, 1976; Braga, 1977; Scherre, 1978; Ponte, 1979; Nina, 1980 e Guy, 1981). De forma geral, todos estes trabalhos apresentam uma conclusão uniforme: a primeira posição do sintagma nominal favorece variavelmente a presença da marca explícita de plural e as demais desfavorecem-na, também variavelmente. A influência da posição tem sido considerada como um dos argumentos decisivos a favor da posição que enfatiza a contribuição africana ao português popular do Brasil, especialmente no trabalho de Guy (1989, p.232-4).

Scherre (1988)³ demonstra que uma análise da concordância nominal que leva em conta apenas a variável *posição linear* encobre regularidades lingüísticas importantes, e propõe uma análise alternativa que considera uma nova variável, advinda do cruzamento entre (1) *posição linear*, (2) *classe nuclear e não nuclear* e (3) *relação* entre classe nuclear e não nuclear. Neste artigo, assumimos a análise proposta por Scherre (1988), acrescida de testes estatísticos complementares, utilizando a versão dos programas VARBRUL de 1988 (Pintzuk, 1988).

Nesta análise alternativa, cujo objetivo central é verificar até que ponto a posição linear é a principal restrição no processo de concordância entre os elementos do SN, são levadas em conta duas variáveis concorrentes: (1) *posição linear* e (2) *posição/classe/relação*.

A variável *posição linear*, subdividida em três categorias, encontra-se exemplificada a seguir:⁴

1) primeira posição:

umas casinha bonitinha/*todos* os anos
suas tias/*novas* escolas/*coisas* lindas
 eles *todos/us* boas ações

2) segunda posição:

umas *casinha* bonitinha/*todos os* anos
 suas *tias/novas* *escolas/coisas* *lindas*
 eles *todos/as* *boas* ações

3) demais posições:

umas casinha *bonitinha*/*todos os* *anos*
 as boas *ações/aquelas* cruzinha *toda*
 uns colega *meu*/três risco *verde*
 as conta quase *toda*/os meus quatro *filhos*
 as três *coisa* mais *importante*
 as mulheres ainda muito mais *antiga*

A variável *posição/classe/relação*, por sua vez, ficou com sete categorias, exemplificadas abaixo:

1) classe não nuclear anteposta na primeira posição (constituída pelos adjetivos, possessivos, quantificadores, artigos, demonstrativos, identificadores e indefinidos que ocorrem na primeira posição do SN):

novas escolas/*suas* tias/*todos* eles
 as boas *ações/aquelas* cruzinha *toda*
determinadas pessoas/*alguns* meses

2) classe não nuclear anteposta na segunda posição (constituída pelos adjetivos, possessivos, quantificadores, artigos, demonstrativos, identificadores e indefinidos antepostos na segunda posição do SN):

todos os anos/*todos* *esses* homens
 os *próprios* vagabundos/*tantas* *outras* famílias
 os *meus* ainda mais velhos amigos/
 as *boas* ações

3) classe nuclear na primeira posição (substantivos e pronomes pessoais que ocorrem na primeira posição do SN):

coisas lindas/ eles todos

4) classe nuclear na segunda posição (substantivos, pronomes pessoais e outras classes gramaticais no lugar do núcleo que ocorrem na segunda posição do SN):

umas *casinha* bonitinha/suas *tias*/
todos *eles*/ trabalha uma, as *outras* fica

5) classe nuclear nas demais posições (substantivos e pronomes pessoais e outras classes gramaticais no lugar do núcleo que ocorrem nas terceira, quarta, quinta e sexta posições do SN):

as boas *ações*/os meus quatro *filhos*/
os meus ainda mais velhos *amigos*/
os melhores *possíveis*

6) classe não nuclear posposta na segunda posição (constituída pelos adjetivos, possessivos, quantificadores, identificadores e indefinidos que ocorrem antepostos na segunda posição do SN):

eles *todo*/eles *mesmo*/
coisas *lindas*/colegas *minha*

7) classe não nuclear posposta nas demais posições (constituída pelos adjetivos, possessivos, quantificadores, identificadores e indefinidos que ocorrem pospostos nas terceira, quarta, quinta e sexta posições do SN):

umas casinha *bonitinha*/
as mulheres ainda muito mais *antiga*/
uns colega *meu*/aquelas cruzinha *toda*/
as conta quase *toda*/uns negócio *qualquer*/
uns vinhos *tais*

Feita a análise quantitativa dos dados assim codificados, os instrumentos estatísticos utilizados revelaram que a variável *posição linear* não tem relevância estatística quando se considera, ao mesmo tempo, a variável

posição/classe/relação. Enquanto, de um conjunto de nove variáveis, a variável *posição/classe/relação* é selecionada como a mais significativa, a variável *posição linear*, além de não ser selecionada, na parte *step up* do programa VARBRUL, é também eliminada na parte *step down*.⁶ Portanto não é apenas a posição linear do elemento no SN que determina a quantidade de marcas formais de plural; os condicionamentos são bem mais complexos do que se pensava anteriormente, conforme nos evidenciam os resultados da variável *posição/classe/relação* apresentados na tabela 1.

Tabela 1
Variável relação/posição/classe

Fatores	Freqüência	Peso relativo
Classe não nuclear na 1ª posição do SN	4065/4167 = 98%	0,87
Classe não nuclear anteposta na 2ª posição do SN	305/317 = 96%	0,86
Classe nuclear na 1ª posição do SN	165/173 = 95%	0,70
Classe nuclear na 2ª posição do SN	2864/5320 = 54%	0,21
Classe nuclear nas demais posições do SN	333/573 = 62%	0,25
Classe não nuclear posposta na 2ª posição do SN	95/139 = 68%	0,26
Classe não nuclear posposta nas demais posições do SN	151/430 = 35%	0,13
Total	7978/11083 = 72%	

Os resultados da tabela 1 podem ser olhados sob três perspectivas: a da posição, a da classe nuclear/não nuclear e a da relação entre a classe nuclear/não nuclear.⁵ Olhando sob a perspectiva da posição, verificamos que a sua influência não é uniforme. A evidência mais forte desta não uniformidade pode ser encontrada na diferença existente entre o efeito da segunda posição nas classes não nucleares antepostas (0,86), nos núcleos (0,21) e nas classes não nucleares pospostas (0,26). Além disso, o efeito das demais posições - terceira, quarta, quinta e sexta - diminui nas classes nucleares pospostas (0,13) mas aumenta nos núcleos (0,25). Conseqüentemente, afirmações generalizantes a respeito do português popular do Brasil com relação ao papel da posição linear que o elemento ocupa no SN carecem de fundamento. Testamos todas as diferenças nos pesos relativos acima e todas elas, exceto a de 0,01, se mostraram significativas, ao nível de 0,05 ou melhor.

Além do que foi dito acima, cumpre observar que a variável cruzada é mais forte do que a variável posição linear isolada. Em análises separadas, a variável *posição/classe/relação* é selecionada em primeiro lugar, dentre um conjunto de nove variáveis, enquanto a variável *posição linear* o é em nono.

Embora tenhamos descartado a influência da *posição linear* como uma influência uniforme, ela ainda se verifica, especialmente nos núcleos da primeira posição. Todavia esta influência deve ser interpretada como uma extensão da influência dos elementos não nucleares antepostos da primeira posição. Como a primeira posição é predominantemente ocupada por classes não nucleares antepostas e como estas classes são muito marcadas (quer na primeira quer na segunda posição), é lícito levantar a hipótese de que seu peso, na primeira posição, seja transferido para qualquer classe que a ocupar. Inclusive, há dialetos do português (interior de Minas Gerais) que apresentam estruturas do tipo "ques" menino levado", em que se verifica a flexão do elemento "que", não flexionável em português.

Se o efeito da posição linear fosse mais forte do que o da posição relativa (anteposição/posposição em relação ao núcleo), seria esperado que a classe não nuclear anteposta na segunda posição fosse pouco marcada, uma vez que os elementos nucleares que ocorrem nesta posição, os mais frequentes (5320 casos de um total de 11083), são muito pouco marcados (0,21). A influência da posição linear não se dá e

o que vemos é a permanência da influência da anteposição em relação ao núcleo.

Concluindo esta seção, podemos verificar que o efeito da posição linear não tem a força que se supunha ter. O que existe é um jogo complexo de influências cruzadas de posição, classe nuclear/não nuclear e relação entre classe nuclear/não nuclear. Conseqüentemente, é temerário utilizar o argumento da posição, nos termos de Guy (1989), como um traço de crioulização advindo de influência da estrutura de línguas africanas.

4. Distribuição geográfica

Um conjunto expressivo de estudos dialetológicos realizados até 1950 e um conjunto igualmente expressivo de estudos sob a perspectiva da teoria da variação lingüística realizados na última década indicam que a variação na concordância nominal em português é um fenômeno geral, independente de localização geográfica.

Com relação aos dialetólogos que abordam estes fenômenos temos os seguintes trabalhos:

- 1) sobre o dialeto caipira (Amaral, 1920);
- 2) sobre o português popular do Nordeste (Monteiro, 1933; Marroquim, 1945);
- 3) sobre o linguajar carioca (Nascentes, 1953) e
- 4) sobre o português de forma geral (Melo, 1946).

Com relação aos estudos realizados sob a perspectiva da teoria da variação lingüística temos o seguinte:

- 1) sobre o português falado no Rio de Janeiro por pessoas de procedência geográfica diversa (Braga & Scherre, 1976);
- 2) sobre o português de Minas Gerais, falado no Triângulo Mineiro (Braga, 1977);
- 3) sobre o português do Rio de Janeiro (Scherre, 1978 e 1988; Guy, 1981);
- 4) sobre o português do Rio Grande do Sul, falado em Porto Alegre (Ponte, 1978);

5) sobre o português do Pará, falado na micro-região Bragantina (Nina, 1980) e

6) sobre o português falado em Brasília. (Dias, 1993).

Da leitura dos diversos trabalhos relacionados acima, pode-se concluir que a variável social determinante das diferenças do comportamento lingüístico é o grau de escolarização e não a procedência geográfica. Se a origem do português do Brasil fosse devida à existência de um pidgin ou de um crioulo de base lexical portuguesa e gramática africana, seria de se esperar que as áreas geográficas que apresentavam maior concentração de escravos em meados do século XIX evidenciassem diferença de comportamento no que toca à concordância nominal e verbal, fato não constatado pelos diversos trabalhos existentes.

A hipótese do papel central da criouliização africana sob foco é ainda facilmente refutada pelo próprio trabalho de Jeroslow (1974, 1975) que deu impulso à discussão quanto às origens crioulas do português popular do Brasil.

Em 1974, Jeroslow fez um estudo sobre o português cearense rural e em 1975 escreveu um artigo sob o título de "Creole characteristics in rural Brazilian Portuguese". É, todavia, fato conhecido na história do Brasil que não houve concentração de escravos de origem africana no Ceará. Em verdade, o Ceará praticamente não teve população negra na época relevante para sustentar a hipótese da africanização do português, seja diretamente das línguas africanas ou, indiretamente, através de um pidgin/crioulo formado pelos escravos africanos.

5. Conclusão

Concluimos, sugerindo uma descrição geral de nosso modelo de desenvolvimento do português popular do Brasil.

1) A língua portuguesa falada em Portugal antes da colonização do Brasil já possuía uma deriva secular que a impulsionava ao longo de um vetor de desenvolvimento.

- 2) No Brasil, este vetor se encontrou com outras forças que ora o reforçavam na direção original, ora o desviavam dessa direção.
- 3) No início uma dessas forças era a pidginização, que exercia uma influência sobre o português através da língua geral tupi e da "língua de preto" europeia, revivificada no Brasil originalmente para uso com os ameríndios. Mais tarde pidgins de base lexical africana também devem ter influenciado, juntamente com outros processos de imitação e de simplificação voluntária.
- 4) Ao longo de toda a história do Brasil o processo de aprendizado teve seus efeitos documentados parcialmente.
- 5) Se existiu uma verdadeira língua crioula de "léxico português e gramática africana", ela cedo se evaporou sem deixar rastros na documentação. Sua possível influência no desenvolvimento do português do Brasil seria indistinguível da de outros pidgins ou crioulos de base não-europeia.

Para completar esse quadro temos que levar em conta a dimensão tempo, ao longo da qual chegaram ao Brasil ondas de populações de diversas origens étnicas. Estes movimentos demográficos tiveram o efeito de reforçar certas forças arroladas acima e de inibir outras. Naturalmente o efeito não foi uniforme nas diversas comunidades lingüísticas e nem mesmo nos indivíduos, mas em termos grosseiros costuma se distinguir três fases de predomínio étnico-demográfico: a fase inicial de predomínio ameríndio, a fase de predomínio africano e por último uma fase importante de reforço europeu e o começo da imigração asiática. Entretanto, a última fase não chegou a esmagar a fase anterior africana com a mesma força que esta teve sobre a fase ameríndia. As primeiras duas fases eram pidginizantes e caracterizadas pelo aprendizado; a última é mais padronizante embora também marcada pelo aprendizado por populações europeias e asiáticas oriundas de terras onde o português era desconhecido.

Através de todas essas fases, segundo nossa visão, o impulso motor do desenvolvimento do português do Brasil veio já embutido na deriva secular da língua de Portugal. Se as sementes trazidas de lá germinaram mais rápido e cresceram mais fortes é porque as condições, aqui, mostraram-se mais propícias devido a uma confluência de motivos.

(Recebido em 22/12/1993)

NOTAS

¹ A escolha deste texto para homenagear Fernando Tarallo foi motivada pelo fato de o assunto a respeito das origens do português do Brasil ser de seu interesse e, ainda mais, por nossa posição ser semelhante a dele, expressa em um texto inédito de 1989, intitulado *On the alleged creole origin of Brazilian Portuguese: untargeted syntactic changes*.

² Seguindo o uso tradicional em estudos crioulos, usamos o termo "base lexical", ou simplesmente "base", para designar a língua da qual provém a maior parte dos itens lexicais usados num pidgin ou num crioulo. Assim, um pidgin ou crioulo de "base portuguesa" tem a maioria de seu vocabulário derivada do português. A estrutura gramatical pode ser independente do vocabulário. De fato, muitos estudiosos são da opinião de que os crioulos de bases europeias (o francês do Haiti, o português de Cabo Verde etc.) possuem uma gramática africana. Quando estamos dizemos que um certo pidgin ou crioulo é de "base x", estamos falando apenas do léxico e não da gramática. Um pidgin ou crioulo de base portuguesa terá necessariamente palavras portuguesas, mas suas estruturas gramaticais poderão ser portuguesas, africanas, ameríndias, ou de qualquer outra origem. Este uso é devido ao fato de que é muito mais fácil identificar as origens do vocabulário do que as fontes da gramática.

³ Os dados analisados por Scherre (1988) são provenientes do banco de dados do Programa de Estudos do Uso da Língua (PEUL), que é um programa de pesquisa desenvolvido desde 1982 por uma equipe de pesquisadores da UFRJ, UFRRJ/CNPq e UFF, através do Departamento de Linguística e Filologia da Faculdade de Letras da UFRJ.

⁴ A variável *posição linear* é comumente subdividida em quatro categorias. Neste caso, optamos por três porque os dados da quarta, quinta e sexta posição são poucos e ficariam muito fragmentados quando do cruzamento entre posição, classe e relação. Além disso, estas posições têm comportamento bastante semelhante.

⁵ A soma dos dados das classes não nucleares da primeira e da segunda posição não perfaz o mesmo total que a soma dos dados das classes nucleares da primeira, segunda e demais posições porque há muitos sintagmas que têm um numeral na primeira ou na segunda posição que, embora faça parte do SN, não constitui dado de nossa análise por não ser flexionável. Só na primeira posição do SN há 1695 numerais (cf. Scherre, 1988, p.236).

⁶ Na parte *step up* o programa trabalha de forma progressiva buscando as variáveis mais importantes do ponto de vista estatístico, ou seja, as que mais dão conta da variação dos dados que estão sendo analisados. Neste caso, o teste começa com as variáveis isoladas e, através de comparações sucessivas, vai

adicionando à análise cada uma das variáveis *selecionadas*, parando quando não mais restar variáveis significativas; na parte *step down* o programa faz exatamente o inverso: parte do conjunto total das variáveis, verificando que variável torna a análise *pior*, eliminando também de forma progressiva tais variáveis (cf. Sankoff, 1988, p.991-2).

REFERÊNCIAS BIBLIOGRÁFICAS

- AMARAL, A. (1920) *O dialeto caipira*. São Paulo, O Livro.
- BORTONI-RICARDO, S. M. (1981) A concordância verbal em português: um estudo de sua significação social. IN: *Ensaio de Lingüística Aplicada ao Português*. Brasília, Thesaurus, p.79-101.
- BRAGA, M. L. (1977) A Concordância de número no sintagma nominal no Triângulo Mineiro. PUC, Rio de Janeiro. Dissertação de Mestrado, inédito.
- BRAGA, M. L. & M.M.P. SCHERRE (1976) A concordância de número no SN na área urbana do Rio de Janeiro. IN: *Encontro Nacional de Lingüística*, 1º, 1976. Anais ... Rio de Janeiro, PUC, p.464-77.
- DIAS, M. C. A. C. (1933) A variação na concordância nominal: um contraste entre o urbano e o rural na fala brasiliense. UnB, Brasília, Dissertação de Mestrado, inédito.
- GRANDGENT, C. H. (1962) *An introduction to Vulgar Latin*. New York, Hafner Pub.Co.
- GUY, G. R. (1981) Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax, and language history. PH.D. Dissertation, University of Pennsylvania, inédito.
- _____ (1989) On the nature and origins of popular Brazilian Portuguese. IN: *Estudios sobre Espanhol de América Y Lingüística Afroamericana*. Bogotá, p.226-44.
- JEROSLOW, E. H. M. (1974) Rural Careense Portuguese: a study of one variety of nonstandard Brazilian speech. Cornell University dissertation.
- _____ (1975) Creole characteristics in rural Brazilian Portuguese. Paper presented at the International Conference on Pidgins and Creoles, University of Hawaii.
- LEITE de VASCONCELOS, J. P. de M. (1970) *Esquisse d'une dialectologie portugaise*. 2ª ed. Lisboa, Centro de Estudos Filológicos.
- NARO, A. J. (1981) The social and structural dimensions of a syntactic change. *Language. LSA*, 57.1:63-98.

- NINA, T. de J. C. (1980) Concordância nominal/verbal do analfabeto na micro-região de Bragantina. PUC, Rio Grande do Sul. Dissertação de Mestrado, inédito.
- MARROQUIM, M. (1945) *A língua do Nordeste - Alagoas e Pernambuco*. 2ª ed. São Paulo, Nacional.
- MELO, G. C. de. (1946) *A língua do Brasil*. Rio de Janeiro, Agir.
- MONTEIRO, C. do R. (1933). *A linguagem dos cantadores*. Tese apresentada à Congregação do Colégio Pedro II. Rio de Janeiro.
- NASCENTES, A. (1953) *O linguajar carioca*. 2ª ed. Rio de Janeiro, Organizações Simões.
- PINTZUK, S. (1988) VARBRUL programs, inédito.
- PONTE, V. M. L. (1979) A concordância nominal de uma comunidade de Porto Alegre. PUC, Rio Grande do Sul. Dissertação de Mestrado, inédito.
- RODRIGUES, J. H. (1983) A vitória da língua portuguesa no Brasil colonial. *Humanidades*. 1.4:21-41. jul/set.
- RODRIGUES, N. (1932) *Os africanos no Brasil*.
- SANKOFF, D. Variable rules. IN: U.AMMON; N. DITTMAR & K.J. MATTHEIER (eds.) *Sociolinguistics - An international handbook of the science of language and society*. Berlin/New York, Walter de Gruyter, 1988, p.984-98.
- SCHERRE, M. M. P. (1978) A regra de concordância de número no sintagma nominal em português. PUC, Rio de Janeiro. Dissertação de Mestrado, inédito.
- _____. (1988) Reanálise da concordância nominal em português. UFRJ, Rio de Janeiro. Tese de Doutorado, inédito.
- SILVA, R. V. M. e. (1992) Português brasileiro - Raízes e trajetórias. *Ciência Hoje*. Rio de Janeiro, 86.15:76-81.
- SILVA NETO, S. da. (1986) *Introdução ao estudo da língua portuguesa no Brasil*. Rio de Janeiro, Presença.
- STURTEVANT, E. H. (1940) *The pronunciation of Greek and Latin*. 2ª ed. Philadelphia, LSA, University of Pennsylvania.
- TARALLO, F. (1989) On the alleged creole origin of Brazilian Portuguese: untargeted syntactic changes, inédito.