

A HIPÓTESE DO SUBSTRATO NO CONTEXTO DA HISTÓRIA SOCIOLINGÜÍSTICA DO PORTUGUÊS POPULAR DO BRASIL

THE SUBSTRATE HYPOTHESIS IN THE
CONTEXT OF THE SOCIOLINGUISTIC HISTORY
OF POPULAR BRAZILIAN PORTUGUESE

DANTE LUCCHESI

Universidade Federal da Bahia/CNPq, Brasil

dante.lucchesi@gmail.com

O artigo discute a transferência de estruturas das línguas indígenas e africanas para a gramática de variedades do português popular brasileiro em sua formação. A hipótese do substrato entra na agenda da pesquisa linguística no Brasil desde que o contato entre línguas passou a ser considerado um fator preponderante na história linguística do país, fortalecendo paralelos entre a origem do português popular brasileiro e das línguas crioulas. O artigo apresenta uma importante formulação substratista na atualidade, baseada no conceito de *relexificação*, e analisa as possibilidades de sua aplicação ao estudo da história sociolinguística do Brasil, considerando sua adequação, fragilidades e potencialidades, dentro da visão adotada aqui de que as variedades populares do português brasileiro passaram, em sua formação histórica, por um processo de transmissão linguística irregular de tipo leve.

Palavras-chave: substrato; português popular do Brasil; história sociolinguística do Brasil; contato entre línguas.

The article discusses the transferency of structures of indigenous and African languages to the grammar of popular varieties of Brazilian Portuguese during its formation. The substrate hypothesis takes place on the linguistic research in Brazil since contact between languages began to be considered a major factor in the linguistic history of the country, strengthening the parallels between the origin of popular Brazilian Portuguese and creole languages. The paper presents one important substratist formulation in creole linguistics, based on the concept of *relexification*, and it analyses its possible application to the study of the sociolinguistic history of Brazil, considering its appropriateness, weaknesses and potential, within the view adopted here that, during their formation, the popular varieties of Brazilian Portuguese underwent a process of irregular linguistic transmission of a light type.

Key words: substrate; Popular Brazilian Portuguese; sociolinguistic history of Brazil; language contact.

1. INTRODUÇÃO

A consideração do contato entre línguas como um fator determinante na história do português popular brasileiro – questão retomada, na década de 1980, por Gregory Guy (1981, 1989) e John Holm (1987, 1992), e desenvolvida posteriormente por Alan Baxter e Dante Lucchesi (Baxter e Lucchesi 1997; Lucchesi 2002, 2004, 2008, 2012a, 2013) – aproxima, em maior ou menor grau, as situações de formação do português popular brasileiro das situações de pidginização e criouliização, dentro de uma perspectiva que traça paralelos nítidos entre o contexto colonial brasileiro e as sociedades de plantação do Caribe e entrepostos comerciais na costa da África (onde vicejaram boa parte das línguas crioulas hoje conhecidas no mundo), em função da importação de cerca de quatro milhões de escravos africanos para o Brasil, no período que se estende, *grasso modo*, de 1550 a 1850. O desenvolvimento dessa perspectiva de análise na agenda da pesquisa linguística no Brasil (Lucchesi *et al.*

2009) coloca necessariamente em questão como as hipóteses formuladas para a gênese das línguas pidgins e crioulas podem contribuir para a compreensão da formação histórica das variedades populares da língua no Brasil.

As duas principais visões sobre a gênese das línguas pidgins e crioulas que historicamente se opõem e disputam a hegemonia no âmbito da Crioulística são a visão dos *universais linguísticos* e a visão da *transferência do substrato*. A ideia de que a gramática das línguas pidgins e crioulas resulta, em grande medida, da combinação dos modelos gramaticais da(s) língua(s) nativa(s) dos falantes adultos do grupo dominado com os itens lexicais da língua do grupo dominante é uma das mais tradicionais na Crioulística (Sylvain 1936; Goodman 1964; Voorhoeve 1973; Koopman 1986; entre outros). Porém, essa visão ficou em segundo plano, nas décadas de 1980 e 1990, com a ascensão no campo da *Hipótese do Bioprograma da Linguagem*, de Derek Bickerton (1981, 1999), que propunha uma explicação para as características estruturais compartilhadas pelas línguas crioulas ao redor do mundo, independentemente da composição das línguas que concorreram para sua formação, com base na ideia de que a regramaticalização da variedade de língua que se formava nas situações de contato linguístico massivo, abrupto e radical ocorria crucialmente quando da sua nativização, em função do dispositivo inato de aquisição da linguagem, com o qual as crianças nascidas na situação de contato reconstituíam a gramática da língua crioula emergente com base na variedade de segunda língua falada pelos adultos do grupo dominado: um pidgin incipiente praticamente desprovido de estrutura gramatical. Contudo, em face dos crescentes questionamentos feitos à hipótese do Bioprograma da Linguagem de Bickerton, a partir do final da década de 1990 (Roberts 2000; Siegel 2008), a hipótese do substrato têm retomado seu espaço no campo da Crioulística, com profícuas teorizações que lançam luzes sobre a formação das línguas pidgins e crioulas, com uma fundamentação empírica bastante consistente (Lefebvre 1997, 1998, 2001; Lumsden 1999a, 1999b; Siegel 2008).

Neste artigo, será focalizada a aplicação da hipótese do substrato à análise da formação histórica das variedades populares do portu-

guês brasileiro. Para cumprir esse objetivo, este artigo se estrutura basicamente em duas seções. Na primeira, será feita uma breve resenha de uma das mais importantes teorizações desenvolvidas no âmbito da crioulistica, dentro da perspectiva do substrato, que se articula em torno do conceito de *relexificação*. Essa formulação será avaliada no contexto das especificidades do caso brasileiro, na segunda e última seção deste artigo. Na conclusão, serão delineadas algumas orientações para a pesquisa que vise a identificar possíveis transferências do substrato na configuração estrutural das variedades populares do português brasileiro.

2. A INFLUÊNCIA DO SUBSTRATO NA FORMAÇÃO DAS LÍNGUAS PIDGINS E CRIOULAS

Uma das teorizações mais recentes desenvolvidas a partir da *hipótese do substrato* baseia-se na ideia de que, na formação da(s) interlíngua(s) que dará(ão) origem à língua pidgin ou crioula, os falantes adultos do(s) grupo(s) dominado(s) combinam a forma fonética dos itens lexicais da língua do grupo dominante (a *língua de superestrato*) com as propriedades semânticas e sintáticas dos itens lexicais correspondentes em suas línguas nativas (as *línguas do substrato*), em um processo denominado *relexificação*. Esse conceito tem sido desenvolvido por Claire Lefebvre (1997, 1998, 2001) e John Lumsden (1999a, 1999b), com base, sobretudo, em suas pesquisas sobre o haitiano, um crioulo de base lexical francesa que é a língua materna da grande maioria da população atual do Haiti. O haitiano se formou em uma sociedade de plantação típica, na qual a população de escravos africanos, em sua maioria falante da língua fongbe – uma língua da família kwa, do oeste africano –, excedia em muito o número de colonizadores franceses, numa proporção de mais de dez africanos para cada francês. Em função disso, Lefebvre e Lumsden afirmam que uma parcela significativa do vocabulário do haitiano é formada por palavras que têm a forma fonética do francês, mas são usadas segundo as propriedades semânticas e sintáticas abstratas da gramática do fongbe.

Lumsden (1999a: 130-133) argumenta que a relexificação é um processo recorrente na aquisição de segunda língua por parte de falantes adultos e apresenta dados de aprendizes franceses do inglês e vice-versa, reproduzidos abaixo:

- 1 *At sixty-five years old they must *retire themselves* because ...
- 2 À soixante-cinq ans, ils doivent *se retirer* en raison ...
- 3 At sixty-five years old they must retire because ...
 - a. ‘Aos sessenta e cinco anos, eles devem se aposentar porque...’
- 4 *Je vais *préparer* pour la fête.
- 5 I will prepare for the party.
- 6 Je vais *me préparer* pour la fête.
 - b. ‘Vou-me preparar para a festa.’

No exemplo (1), produzido por um aprendiz francês do inglês, o verbo *retire* é usado de acordo com as especificações gramaticais do verbo correspondente em francês, segundo as quais ele seria um verbo pronominal – cf. exemplo (2) –, mas em desacordo com as propriedades gramaticais desse verbo em inglês – cf. exemplo (3). O contrário acontece na frase em 4, produzida por um aprendiz inglês do francês, no qual falta o clítico – cf. variante padrão do francês apresentada em (6), porque o verbo correspondente em inglês não é pronominal – cf. exemplo (5). Esses são exemplos de relexificação, em que o item lexical da língua alvo é empregado segundo as especificações gramaticais da língua materna do aprendiz adulto de uma segunda língua. Entretanto, nos processos de aquisição sistemática de segunda língua, no qual o aprendiz é treinado para assimilar os modelos gramaticais da língua alvo, esses erros são corrigidos, e a relexificação é eliminada, quando o aprendiz atinge sua meta, ou seja, adquirir uma competência linguística semelhante à dos falantes nativos da língua alvo.

Situação muito distinta é aquela em que se formam as línguas pidgins e crioulas, que também têm como ponto de partida a aquisição de uma segunda língua por parte de uma grande massa de falantes adultos (Wekker 1996; Field 2004; Siegel 2008). A grande maioria das línguas pidgins e crioulas se formaram em contexto de escravidão ou similares (situações outras de trabalho forçado “contratado”), em que o número de falantes dos grupos dominados era muito maior do que os falantes do grupo dominante. Nesse contexto, o acesso que os falantes das outras línguas tinham aos modelos gramaticais da língua dominante era muito restrito. Na verdade, a comunicação inicial entre dominantes e dominados era feita através de um código de comunicação emergencial, denominado *jargão* ou *pré-pidgin* e formado por um reduzido elenco de itens lexicais da língua de superestrato. O pré-pidgin seria empregado basicamente para preencher as funções comunicativas básicas, tais como: dar uma ordem ou transmitir informação simples (Siegel 2008). Portanto, para além do restrito acesso aos modelos da língua dominante, na pidginização e/ou criouliização, os falantes das outras línguas também não tinham como objetivo, na maioria das vezes, adquirir uma competência linguística semelhante à dos falantes da língua de superestrato (Rampton 1997). Seu objetivo era somente dispor de código para se comunicar com os indivíduos do grupo dominante e, sobretudo com os demais indivíduos do grupo dominado, já que a maioria das línguas pidgins e crioulas emergiu no seio de comunidades pluriétnicas e plurilíngues, formadas à força, em função da escravidão e outras formas de dominação.

A reestruturação do jargão ou pré-pidgin, que torna possível o surgimento de uma língua pidgin ou crioula, depende crucialmente de sua utilização para a comunicação entre os indivíduos dos grupos dominados (Whinnom 1971). Quando isso ocorre, amplia-se sobremaneira o espectro funcional do pré-pidgin, em razão das relações de cooperação e solidariedade que formam entre esses indivíduos, base para ampliação das funções comunicativas em sua interação verbal. Com um limitado acesso aos modelos da língua de superestrato e, muitas vezes, sem ter a preocupação de reproduzir fielmente esses modelos, os falantes adultos das línguas do substrato vão desenvol-

vendo o restrito inventário lexical da língua dominante com base nos dispositivos gramaticais de suas línguas nativas. Esse contexto é altamente favorecedor, não apenas da ocorrência de processos de relexificação, mas também da cristalização desses processos na sedimentação de soluções que formarão a gramática na língua pidgin ou crioula emergente.

A expansão funcional e consequente (re)gramaticalização do jargão ou pré-pidgin entre os falantes adultos dá origem ao que se denomina *pidgin expandido*, variedade linguística que se torna o principal veículo de comunicação verbal da nova comunidade de fala plurilíngue, não obstante se conserve, por várias gerações, como segunda língua da maioria dos seus falantes. Tal é caso do tok pisin, em Papua Nova Guiné (Mühlhäusler 1986), que tem sido usado como língua de comunicação interétnica em um país com várias centenas de línguas nativas, desde finais do século XVIII; só se nativizando, de forma representativa, a partir da década de 1970, não obstante já fosse usado nos diversos níveis sociais de interação verbal, inclusive na modalidade escrita e nos meios institucionais, no parlamento e no rádio e televisão.

À medida que essa variedade de segunda língua mais ou menos estruturada gramaticalmente vai-se tornando a língua materna dos indivíduos que nascem nessa nova comunidade, vai ocorrendo a emergência da língua crioula. Entretanto, a crioulaização pode ocorrer em qualquer ponto do *continuum* de pidginização, desde o pré-pidgin até o pidgin expandido.¹ No caso das línguas crioulas da costa da África e do Caribe, não se pode afirmar com precisão em que ponto do *continuum* de pidginização a crioulaização ocorreu, pois os únicos dados empíricos disponíveis atualmente provêm das análises dessas línguas crioulas que se fizeram nas últimas décadas, complementadas por análise mais recuadas no tempo, a partir de registros históricos esparsos. No geral, não se dispõe de registros históricos das variedades de segunda língua faladas pelos escravos

1 Segundo Bickerton (1999), só se poderia falar em crioulaização propriamente dita, se a nativização ocorresse na fase inicial da pidginização, quando as crianças tivessem como modelo de aquisição de sua língua materna um pré-pidgin gramaticalmente incipiente.

africanos que deram origem a essas línguas crioulas. Portanto, a base empírica para as hipóteses do substrato decorrem basicamente da comparação atual de estruturas gramaticais da língua crioula com estruturas igualmente atuais da(s) língua(s) que supostamente teria(m) sido hegemônica(s) entre os adultos que formaram inicialmente a comunidade crioulófona.

As teorizações sobre como ocorre a transferência do substrato se baseiam nas correspondências observadas por esse método de análise, não obstante os problemas empíricos que tal opção enseja. Tais problemas serão comentados adiante. Antes, será feita a apresentação das formulações de Claire Lefebvre e John Lumsden sobre a formação do crioulo haitiano, que giram em trono do conceito de *relexificação*. De acordo com Lefebvre (2001) e Lumsden (1999a, 1999b), a *relexificação* desempenharia um papel central na gramaticalização do pidgin, ao lado dos processos de *reanálise* e *nivelamento dialetal*. Nessa perspectiva, o léxico das línguas pidgins e crioulas seria formado da seguinte maneira:

As entradas lexicais dos léxicos das línguas de substrato são copiadas, e as representações fonológicas nessas entradas lexicais copiadas são substituídas por representações fonológicas derivadas de formas fonéticas da língua do superestrato ou por formas nulas. Esse segundo passo é denominado *re Etiquetagem* (*relabel*). A escolha da forma fonética pertinente da língua de superestrato para reetiquetar a entrada lexical copiada é baseada em seu uso em contextos semânticos e pragmáticos específicos, de modo que a semântica da forma do superestrato deve ter alguma coisa em comum com a semântica da entrada lexical que está sendo reetiquetada². (Lefebvre 2001: 11)

Dessa forma, o resultado da relexificação é uma unidade lexical com uma forma fonética derivada da língua do grupo dominante com as especificações semânticas e sintáticas de uma das línguas do substrato (Lumsden 1999a; Lefebvre 2001). Como exemplos de *relexificação* de *itens lexicais* (ou seja, de *palavras referenciais*),

2 Traduzido do original inglês.

em pidgins expandidos, temos o caso do verbo *harim* do tok pisin, que deriva do verbo inglês *hear* ‘ouvir’, mas exhibe as propriedades semânticas do verbo *igub*, da língua usan, uma das línguas do substrato do tok pisin. Assim como verbo *igub* em usan, o verbo *harim* pode ser empregado em tok pisin, tanto com o significado de ‘ouvir’, quanto de ‘cheirar’ (Reesink 1990: 290). Outro exemplo pode ser encontrado no pidgin das Ilhas Salomão, também na região da Melanésia, no qual o verbo *sidaon* (derivado do inglês *sit down*) é usado pelos falantes mais velhos do pidgin salomão de uma maneira absolutamente equivalente àquela do verbo *to’oru* da língua kwaio do substrato, cobrindo o espectro de significação dos verbos ‘sentar’, ‘viver’ e ‘esperar’ (Keesing 1988: 239). Já o problema da relexificação das categorias funcionais seria mais complexo e igualmente mais relevante, no processo de pidginização/crioulização.

Há um consenso entre os estudiosos do contato entre línguas de que variedades linguísticas como o jargão e o pré-pidgin não possuem categorias funcionais, pelo menos realizadas foneticamente. Uma das razões para isso seria que os falantes adultos das outras línguas não perceberiam as partículas gramaticais da língua alvo, porque, no geral, essas partículas têm uma forma fonética reduzida e átona. Além disso, os próprios falantes da língua alvo suprimiriam essas categorias gramaticais em suas produções verbais, na interação com os falantes das outras línguas, como previsto na teoria do *foreigner talk*. Lumsden (1999a: 142-143) argumenta que os usuários do jargão ou pré-pidgin formariam as frases utilizando o restrito inventário de itens lexicais extraídos da língua alvo, com base nas categorias funcionais de suas línguas nativas, só que essas categorias funcionais não teriam expressão fonética. Assim, seriam produzidas frases com estrutura superficial idêntica, mas com representações estruturais subjacentes diferenciadas, consoante a gramática da língua nativa de cada usuário. Esse fato explicaria a intercompreensão entre falantes de línguas tipologicamente diferenciadas.

Porém, diferentemente do jargão ou pré-pidgin, os pidgins expandidos e os crioulos possuem categorias funcionais realizadas foneticamente, mesmo que categorias funcionais realizadas foneticamente na língua de superestrato (como a concordância nominal e

verbal) não se conservem no crioulo ou no pidgin dela derivado. O que interessa aqui é que, no geral, as categorias funcionais realizadas no pidgin expandido ou no crioulo se diferenciam qualitativamente, em suas propriedades sintáticas e semânticas, da categoria funcional correspondente na língua de superestrato. Considerando os elementos disponíveis no jargão ou pré-pidgin, que é base para o processo de regramaticalização do pidgin ou crioulo, haveria basicamente dois caminhos a serem seguidos pelos agentes da pidginização/crioulização. Um mais evidente seria a fixação de categorias funcionais foneticamente nulas, o que fundamenta a ideia de que os pidgins e crioulos seriam línguas morfologicamente mais simples do que suas línguas lexificadoras (McWhorter 1998, 2001; Siegel 2008).

Outro caminho seria o emprego da forma fonética de palavras lexicais para desempenhar a funções gramaticais, através do processo de *reanálise*, que Lumsden (1999b: 150) define como a associação da forma fonológica de uma palavra lexical com a entrada lexical de uma categoria funcional no interior de uma mesma língua³. É muito frequente a conversão de advérbios ou expressões adverbiais em marcadores de tempo, modo ou aspecto, através da reanálise. Tal é o caso da expressão *by and by* do inglês, da qual deriva a forma fonética do marcador de futuro em tok pisin *bai*, através do advérbio *baimbai* desse mesmo pidgin. Outro exemplo é marcador de futuro *lo* do crioulo português de Malaca, no sudeste asiático, que deriva do advérbio *logo*, do português (Baxter 1988).

Haveria, contudo, uma terceira via, que também se tem mostrado bastante produtiva. Trata-se da utilização de palavras gramaticais que têm uma forma fonética forte para fornecer a forma fonética das categorias funcionais do pidgin ou crioulo emergente. O que se observa, entretanto, nesses casos, é que essa forma fonética forte é associada à entrada lexical de outra categoria funcional, através do processo de *relexificação* aqui referido. Tal é o caso da partícula locativa do francês *là*, que funciona como artigo definido no crioulo

3 É óbvio que deve haver alguma interseção semântica entre a palavra lexical e a categoria funcional para que a reanálise ocorra.

francês do Haiti, bem como da forma foneticamente forte do pronome de 3ª pessoa do plural *eux*, que também é empregado como pluralizador nominal nessa mesma língua. Vamos examinar agora esses dois casos de relexificação.

Segundo Lefebvre (2001: 14-20), o artigo definido *la* do haitiano teria sido formado por meio da relexificação, pois há um nítido paralelo de suas propriedades gramaticais com o artigo definido *ó* do fongbe, língua africana hegemônica no substrato do haitiano. Lefebvre argumenta que o comportamento desse determinante no fongbe e no haitiano se diferencia qualitativamente do artigo definido do francês, a língua lexificadora deste último. Tanto o determinante *la* (e seus alomorfes *a*, *an*, *nan* e *lan*) do haitiano, quanto o *ó* (e seu alomorfe *ón*) do fongbe exibem as seguintes características: (i) são pós-nominais; (ii) não se flexionam em gênero e número; (iii) têm um valor anafórico obrigatório, indicando que a informação veiculada pelo SN em que figuram já é do conhecimento dos participantes da conversação – cf. exemplo (7); (iv) não podem ser usados com nomes de massa e SN de referência genérica – cf. exemplo (8); e (v) aceitam a interposição de uma oração relativa entre ele e o núcleo do SN, cf. exemplo (9):

7 *liv la* (haitiano)

vî ó (fongbe)

livro DEF

‘o livro (do qual estamos falando)’

8 *Pen (*la) Bon pou lasante.* (haitiano)

*Wòxúnú (*ó) nyón nu làn mèyén.* (fongbe)

pão (*o) bom para saúde

‘(*o) pão é bom pra saúde’

9 *Mounn ki pati a.* (haitiano)

Súnù dé é yì ó. (fongbe)

homem que saiu o

‘O homem que saiu.’

Em francês, o artigo definido se flexiona em gênero e número (*le, la, les*), pode ser usado com nomes de massa e de referência genérica e não aceita a interposição de uma oração relativa:

10 *Le pain est bom pour la santé.*

O pão é bom para a saúde.

11 **Le qui est parti homme*

*O que saiu homem

Portanto, Lefebvre conclui que as propriedades gramaticais do artigo definido do haitiano não são derivadas do francês, mas do fongbe. Mais do que isso, Lefebvre (2001: 16-17) argumenta no sentido de que a forma fonética do artigo definido do haitiano não deriva da forma do feminino do artigo definido francês. O ponto de partida de Lefebvre é a forma fonética de muitas palavras do haitiano que incorporam o artigo definido do francês, como se na tabela abaixo:

Haitiano	SIGNIFICADO	Francês	SIGNIFICADO
larivyè	rio	la rivière	o rio
listwa	história	l’histoire	a história
latè	terra	la terre	a terra
zwazo	pássaro	les oiseaux	os pássaros

Tabela 1: palavras do haitiano que incorporaram o artigo definido francês em sua formação.

O significativo número de palavras do haitiano que têm essa composição revela que os criadores do crioulo haitiano não percebiam o artigo definido em francês, foneticamente aglutinado ao nome. O emprego do artigo definido, em sua forma atual, junto a esses nomes – *larivè a* ‘o rio (do qual estamos falando)’ – atesta que o conjunto original DET + NOME do francês se reduziu apenas à forma fonética do nome em haitiano. Lefebvre (2001: 18-20) argumenta, então, que a partícula locativa ou enfática *là*, que pode figurar no final do SN em francês, seria a fonte da forma fonética do artigo definido do haitiano:

12 *Cet homm-lá vient d'arriver.* (francês)

Esse/aquele homem que acabou de chegar.

13 *L'homm-lá vient d'arriver.* (francês popular)

O homem (!) que acabou de chegar.

Obviamente que as propriedades gramaticais da partícula locativa/enfática *là* do francês divergem em grande medida das propriedades gramaticais do artigo definido do haitiano. Teria sido apenas a distribuição similar dessa partícula com a do artigo definido do fongbe (pós-nominal) o que levou os criadores do pidgin/crioulo a tomá-la para reetiquetar a entrada lexical abstrata do artigo definido dentro dos parâmetros da gramática do fongbe.

Os dêiticos seriam outras palavras funcionais do haitiano que teriam sido criadas por meio da relexificação (Lefebvre 2001: 20-28). O haitiano possui apenas dois dêiticos, *sa* e *sila*, que podem ser usados, tanto com nomes de seres animados, quanto com nomes de seres inanimados, e são pós-nominais. *Sila* só é usado para referir objetos que estão distante do falante, contendo o traço semântico [+distante], enquanto *sa* não tem valor marcado para distância, sendo empregado, tanto com objetos próximos, quanto distantes do falante, como se pode ver nos seguintes exemplos:

14 *mounn sa* // *bag sa*

esse/aquele homem // esse/aquele anel

- 15 *mounn sila* // *bag sila*
aquele homem // aquele anel

O fongbe tem um sistema idêntico, com apenas dois dêiticos, *éló* e *éné*, que têm as mesmas características de *sá* e *sila*, como se pode ver abaixo:

- 16 *sínù éló* // *àlòké éló*
esse/aquele homem // esse/aquele anel
- 17 *sínù éné* // *àlòké éné*
aquele homem // aquele anel

Esses dêiticos do haitiano e do fongbe podem coocorrer, da mesma maneira, com possessivo, artigo definido e marcador de plural, como se vê na tabela abaixo:

<i>esse</i>	
Haitiano	Fongbe
bag mwen sa a yo	àlòké ce éló ó lè
anel POSS DEIT DEF PL	anel POSS DEIT DEF PL
esses meus anéis	esses meus anéis
<i>aquele</i>	
Haitiano	Fongbe
bag mwen sila a yo	àlòké ce éné ó lè
anel POSS [+distante] DEF PL	anel POSS [+distante] DEF PL
aqueles meus anéis	aqueles meus anéis

Tabela 2: propriedades semânticas e sintáticas dos dêiticos locativos em haitiano e fongbe.

Assim, “as entradas lexicais do haitiano são diferentes das do fongbe somente em suas representações fonológicas” (Lefebvre 2001: 26). A relexificação teria ocorrido, então, da seguinte maneira:

A entrada lexical copiada que corresponde ao dêitico geral *éló* foi reetiquetada com base na matriz fonética do pronome dêitico geral francês *ça* [sa], resultando no dêitico geral haitiano *sa*. Da mesma forma, a entrada lexical copiada que corresponde ao dêitico fongbe [+distante] *éné* foi reetiquetada com base na forma fonética *cela* ou *celui-là*, resultando no pronome demonstrativo [+distante] haitiano *sila*⁴. (Lefebvre 2001: 26-27)

Já a realização do caso genitivo em haitiano constitui uma situação em que, segundo Lefebvre (2001: 28-29), a entrada lexical da categoria funcional da língua do substrato assume uma forma fonética nula no crioulo, pois, enquanto o fongbe tem o morfema de genitivo pós-nominal *tón* – cf. exemplo (18) –, e o francês marca esse caso com a preposição *de* ou *à* – cf. exemplo (19) –, no haitiano o genitivo é expresso por uma estrutura de justaposição – cf. exemplo (20):

18 *távò Kòkú tón ó*

mesa Koku GEN DEF

a mesa de koku

19 *la table de/à Jean*

a mesa de João

20 *tab JanØ na*

mesa Jan DEF

a mesa de João

4 Traduzido do original inglês.

Lefebvre e Lumsden (1992) assumem que há uma categoria funcional de caso nula em haitiano, com base no princípio de que a atribuição de caso é requerida pela Gramática Universal (Chomsky 1981). Assim sendo, lançam a hipótese de que “a entrada lexical associada com fongbe *tón* foi copiada, mas não reetiquetada”.

Por fim, a formação do pluralizador nominal haitiano *yo* seria mais complexa combinando a relexificação com o nivelamento dialetal (Lefebvre 2001: 30-34). Esse pluralizador pós-nominal tem um paralelo nítido com a partícula *lè* do fongbe, pois sua presença implica sempre uma interpretação [+definido] do SN, o que não impede que ele coocorra com o artigo definido (nesse caso, o pluralizador sempre segue o artigo) – cf. exemplo (21). Já os SNs com valor [-definido] não têm marca de plural – cf. exemplo (22):

21 *krab yo / krab la yo* (**krab yo la*) [haitiano]

àsón lè / àsón ó lè (**àsón lè ó*) [fongbe]

os carangueijos

22 *M'achte krab.* [haitiano]

N'xó àsón. [fongbe]

Eu comprei caranguejos.

A marcação de plural nessas duas línguas difere claramente do que ocorre em francês, em que o plural é marcado nos determinantes pré-nominais, devendo ocorrer também nos nomes indefinidos, como se vê abaixo:

23 *Jean a mangé des pommes.*

João comeu maçãs.

Há, contudo, uma diferença entre o fongbe e o haitiano, pois no haitiano o *yo* é o pronome de 3ª pessoa de plural – *yo pati* ‘eles saí-

ram' –, que funciona também como pluralizador, enquanto no fongbe o pronome da 3ª pessoa do plural tem uma forma distinta: *yè yì* 'eles saíram'. A forma fonética da partícula do haitiano seria derivada da forma forte do pronome da 3ª pessoa do plural *eux* do francês. A extensão do uso da forma do pronome pessoal para o de pluralizador nominal encontraria paralelo, não no fongbe, mas em várias outras línguas da família kwa, do oeste-africano, como o ewe, o yoruba, o mandinga e o edô, nas quais uma mesma partícula desempenha as duas funções. Lefebvre argumenta que na formação inicial do crioulo haitiano, haveria pelo menos dois dialetos, um do tipo fongbe, com duas entradas lexicais distintas (uma para o pronome pessoal, outra para o pluralizador), e um do tipo ewe, com apenas uma entrada lexical para as duas funções. Teria ocorrido então o processo de nivelamento dialetal, no qual teria predominado o dialeto com apenas uma entrada lexical, o que explica as atuais características da partícula *yo* do haitiano.

O que foi dito até aqui pode fornecer a base para a seguinte sistematização. De acordo com teorizações recentes que tomam como ponto de partida a *hipótese do substrato*, os processos de *relexificação* e *reanálise* desempenhariam um papel central na regramatização do jargão ou pré-pidgin que conduz à formação das línguas pidgins e crioulas. Nos casos em que há uma superposição clara entre a estrutura gramatical da língua crioula e a da língua do substrato em oposição à forma da mesma estrutura gramatical na língua do superestrato, como no caso da coincidência estrutural na forma do artigo definido e dos demonstrativos entre o haitiano e o fongbe, em oposição ao francês (cf. supra), a hipótese do substrato parece ser a mais adequada, independentemente da formalização teórica utilizada para fundamentar essa hipótese⁵. Porém, nos casos de simplificação morfológica, como se observa na marcação do caso genitivo em haitiano, a hipótese do substrato, fundamentada na formulação de *cópia da entrada lexical sem reetiquetagem* (cf. supra), não seria a mais adequada. Nesses casos, seria mais plausível assumir um

5 Além do conceito de *relexificação*, aqui apresentado, há atualmente outras formulações concorrentes entre os substratistas, como o conceito de *transferência*, formulado por Siegel (2008), que não será tratado neste artigo.

processo geral de simplificação morfológica que tem sua origem na aquisição maciça de uma segunda língua por parte de falantes adultos com acesso restrito aos modelos gramaticais da língua alvo, independentemente das características gramaticais das línguas nativas desses falantes. A conjugação de estudos de aquisição de segunda língua por adultos com análises de línguas pidgins e crioulas tem fornecido bases empíricas e teóricas bem consistentes para essa visão, que exclui a transferência do substrato (Shumann 1978; Klein e Perdue 1997; Wekker 1996). Assim, a hipótese que se fundamenta nas restrições de tal processo de aquisição de segunda língua é epistemologicamente mais adequada porque é mais econômica, ou seja, recorre a menos construtos teóricos, de modo que a formulação da relexificação, ou qualquer outra que busque estabelecer uma relação com o substrato nesses casos, torna-se desnecessariamente excessiva (Lucchesi 2012b: 269)⁶.

No que concerne ao tema deste artigo, a questão que se coloca é: teria havido processos de mudança induzidos pelo contato entre línguas, na formação de pelo menos algumas das variedades populares do português brasileiro, que se aproximam dos processos de mudança que se enquadram nesse esquema explicativo das transferências de substrato, ou, ao contrário, teriam ocorrido apenas processos de simplificação morfológica que se desenrolam independentemente das influências do substrato? A segunda opção parece ser, em princípio, a mais provável em face da menor radicalidade das situações de contato presentes na história sociolinguística do Brasil *vis-à-vis* as situações prototípicas de pidginização e crioulização (Lucchesi 2008 e 2009).

Além dessa questão teórica mais geral, muitas das restrições que se levantam contra a hipótese do substrato, como a confiabilidade da identidade do alegado substrato de certas línguas crioulas, se colocam no caso brasileiro, de forma, inclusive, mais crítica. Algumas dessas questões serão abordadas na próxima seção.

6 Essa questão será retomada na próxima seção, no contexto da formação das variedades populares do português brasileiro.

3. A HIPÓTESE DO SUBSTRATO NO CONTEXTO SÓCIO-HISTÓRICO BRASILEIRO

Uma das críticas mais contundentes feitas à hipótese do substrato é confiabilidade da identificação das línguas de substrato que efetivamente contribuíram para a formação da língua pidgin ou crioula em foco. A ausência de registros históricos fidedignos constitui, com efeito, um sério obstáculo para a hipótese do substrato, já que boa parte das línguas crioulas hoje reconhecidas como tal se formou no contexto da escravidão de populações africanas. Grande parte do tráfico negreiro e das situações de escravidão ocorreram na clandestinidade e *a fortiori* são escassamente documentadas. Portanto, é muitas vezes difícil identificar que povos teriam constituído efetivamente o substrato de uma determinada língua crioula. Ocorre também que a rota de abastecimento de escravos podia mudar, alterando completamente a composição étnica do substrato. A esse respeito, Arends (1994: 18) fornece um exemplo muito significativo acerca da formação dos crioulos do Suriname, tais como o sranan e o saramacan. A ideia geral era que, durante todo o período do tráfico negreiro promovido pela colonização holandesa (de 1741 a 1815), os escravos teriam sido trazidos da região que vai de Serra Leoa à Costa do Marfim. Com isso, as principais línguas do substrato seriam as línguas mandê e wolof. Porém, dados mais recentes informam que essa região só desempenhou um papel relevante a partir de 1740, quando o sranan e o saramacan supostamente já estariam formados. No período da sua formação (aproximadamente entre 1650 e 1720), a região que mais supriu o Suriname de escravos teria sido a *Costa dos Escravos*, que compreende hoje o Togo e o Benin, e, de outra parte, a região que compreende o Gabão, o Congo, o Zaire e Angola, região de línguas bantas. Portanto, ao invés de uma suposta influência mandê e wolof, as influências do substrato nos crioulos do Suriname deveriam ser buscadas nas línguas fongbe, ou nas línguas bantas, como o quicongo e o quimbundo.

Esse problema assume proporções ainda mais dramáticas no contexto brasileiro, que reúne, em seu substrato, mais de mil línguas autóctones, faladas por pelo menos um milhão e meio de índios quan-

do se inicia a colonização portuguesa, às quais se agregam muitas dezenas de línguas africanas, faladas por cerca de quatro milhões de indivíduos, trazidos pelo tráfico negreiro, entre 1550 e 1850, aproximadamente. Esse imenso mosaico de multilinguismo se distribui por um território de dimensões continentais, através de diversas composições sociolinguísticas bem mais complexas e diferenciadas do que as sociedades de plantação do Caribe, em sua maioria insulares, onde vicejaram as línguas crioulas, com forte influência do substrato africano.

Para além da dificuldade em identificar o substrato, em face da escassez de registros históricos sobre a composição etnolinguística da sociedade brasileira, a hipótese da influência do substrato tem de lidar com as possíveis clivagens na formação histórica das atuais variedades do português brasileiro, separando as áreas em que predominaram as populações indígenas, nomeadamente, na Amazônia e na região do Planalto Central, das regiões com uma presença africana mais maciça, como a Bahia, Pernambuco e o Maranhão, no Nordeste. Não há qualquer relação filogenética entre as línguas indígenas brasileiras e as línguas africanas trazidas para o Brasil. Assim, qualquer característica linguística supostamente derivada de um substrato africano, por exemplo, não deveria estar presente em áreas de substrato indígena, a não ser que se demonstrasse que, por uma improvável coincidência, essa característica estivesse presente também na(s) língua(s) indígena(s) daquela região. O contrário também se aplica, hipóteses de influência do substrato indígena deveriam em princípio delinear uma isoglossa, circunscrevendo a variedade de português formada naquela região onde predominavam os falantes daquele povo indígena.

Porém, há razões suficientes para pensar que uma influência do substrato africano seja mais viável do que a das línguas indígenas. Excetuando-se o primeiro século da colonização, os contatos com as populações indígenas sempre ocorreram nas franjas da sociedade brasileira, e o genocídio das populações autóctones fez com que decrescesse drasticamente a representatividade demográfica do elemento indígena na composição da sociedade brasileira, já nos primeiros séculos da colonização europeia. Some-se a isso a enorme diversida-

de das línguas indígenas brasileiras, separadas em pelo menos cinco famílias linguísticas tipologicamente diferenciadas. O elemento africano, demograficamente mais representativo, integrou-se de forma muito mais ampla e profunda na sociedade brasileira, exercendo um protagonismo linguístico bem mais significativo do que o do segmento indígena (Mendonça 1933; Ribeiro 1995; Lucchesi 2009).

No que concerne à identificação etnolingüística da população do substrato africano na história do Brasil, existe uma divisão consagrada pela tradição historiográfica, na qual a Bahia e o Maranhão se destacam como áreas de substrato kwa, em contraste com restante do Brasil, nomeadamente o Rio de Janeiro e Pernambuco, onde teriam predominado os escravos bantos, com hegemonia dos falantes do quimbundo, mas também com uma presença significativa das línguas quicongo e umbundo (Rodrigues [1933] 2004; Matoso 2003; Castro 2002). Na Bahia, a partir do século XVIII, os falantes de iorubá (aqui chamado *nagô*) seriam maioria no substrato africano, seguidos dos falantes da língua ewe (aqui chamada *jeje*). Já o Maranhão seria área predominantemente jeje, isto é ewe. Minas Gerais seria, em princípio, uma zona de predomínio banto, como o atesta a comunidade de Tabatinga, que ainda usa um código restrito de comunicação intragrupal, composto de um vocabulário quimbundo (Queiroz 1998). Entretanto, há um importante documento histórico, um manual de conversação da língua fon, escrito em Vila Rica, em 1734, que aponta para uma igualmente representativa presença de falantes dessa língua kwa na região (Castro 2002). De qualquer sorte, o dialeto mineiro, por suas especificidades, seria um forte candidato a apresentar alterações estruturais mais profundas.

Nesse cenário, uma via de investigação a ser perseguida seria a de identificar características linguísticas que estivessem mais presentes na Bahia e/ou no Maranhão, como possíveis candidatos a uma eventual transferência do substrato kwa. Ou o contrário, apresentar características linguísticas do português popular de estados, como o Rio de Janeiro e Pernambuco, derivadas do substrato banto, que estariam menos presentes na Bahia e no Maranhão. Nessa perspectiva, o acervo do Projeto ALIB, do Atlas Linguístico do Brasil, distribuído por uma rede de 252 pontos que cobrem todo o território

brasileiro (Cardoso e Mota, 2012), constitui uma valiosa fonte de informação empírica para a verificação de hipóteses acerca de influências do substrato.

Contudo, a ideia corrente de uma homogeneidade do português popular brasileiro militaria contra eventuais influências do substrato. A comprovação de uma transferência do substrato para o português popular deveria, pelo menos em parte, desconstruir essa visão de uma generalizada homogeneidade. É bem verdade que a história do Brasil é marcada por maciços deslocamentos populacionais, particularmente da população escrava africana (Lucchesi 2009). No século XVIII, há um enorme afluxo de escravos dos decadentes engenhos do Nordeste para Minas Gerais, na verdadeira corrida do ouro e das pedras preciosas em curso na época. Já no século seguinte, o deslocamento se dá das exauridas minas de ouro e pedras preciosas para as emergentes plantações de café do Vale do Paraíba, no Rio de Janeiro, e do planalto paulista. E a população do planalto paulista, por sua vez, teve uma importante participação na colonização da região sul, particularmente o território do Estado do Paraná. Já o êxodo rural, ocorrido ao longo do século XX, com seu pico entre as décadas de 1960 e 1990, levou uma grande massa de nordestinos para a região sudeste, sobretudo para periferia das duas grandes metrópoles do país, Rio de Janeiro e São Paulo. Verifica-se igualmente uma forte corrente migratória do Nordeste para a região amazônica, desde meados do século XIX até os dias atuais. Por fim, o *boom* do agronegócio no centro-oeste do país, ocorrido desde as últimas décadas do século passado, atraiu para essa região expressivos contingentes de sulistas.

Esses sucessivos deslocamentos podem explicar a suposta homogeneidade diatópica do português popular brasileiro, em função do nivelamento linguístico que necessariamente se teria implementado no bojo de tais processos, o que teria diluído eventuais transferências do substrato ocorridas no passado. Caberia aos postulantes de influências do substrato demarcar isoglossas até então desconhecidas que atestem a transferência de alguma característica estrutural de língua africana, ou mais remotamente indígena; ou demonstrar, ao contrário, a difusão de uma característica do substrato, em função

dos deslocamentos populacionais, ao invés de sua diluição; o que seria, em princípio, muito difícil de fazer com fundamentação empírica consistente, devido à amplitude e complexidade das variáveis envolvidas.

Mas talvez haja uma questão que preceda a discussão sobre a distribuição geográfica das influências do substrato. O que estaria, então, em jogo seria a própria ocorrência das transferências do substrato nos contextos sociolinguísticos que teriam predominado na formação histórica das atuais variedades linguísticas brasileiras.

Dentro da visão que se baseia no conceito de *transmissão linguística irregular de tipo leve*, a formação das variedades populares do português brasileiro seria significativamente distinta da formação das línguas pidgins e crioulas prototípicas (Lucchesi 2008, 2009 e 2012a). Dentro dessa visão, se ocorreram processos radicais de pidginização e crioulização no Brasil, esses processos teriam sido efêmeros e localizados, concentrados principalmente no século XVII, quando a população de toda a sociedade brasileira não chegava sequer a 300 mil habitantes. Dessa forma, a grande maioria das variedades do português brasileiro teria se formado no contexto de uma *transmissão linguística irregular de tipo leve* que se diferenciaria dos processos típicos de pidginização/crioulização da seguinte maneira:

a originalidade da estrutura gramatical das línguas pidgins e crioulas decorre da impossibilidade de os seus falantes assimilarem os elementos gramaticais da língua alvo, devido à situação de exploração, repressão e segregação a que são submetidos.

Por outro lado, o aumento do acesso dos falantes das outras línguas e/ou de seus descendentes aos modelos gramaticais da língua alvo inibe os processos de transferência de substrato e de gramaticalização, cruciais para a formação da estrutura gramatical das línguas pidgins e crioulas, qualitativamente distinta da estrutura gramatical da língua alvo. Por gozar de maior prestígio, as variantes gramaticais da língua do grupo dominante acabam por prevalecer sobre as estruturas das línguas do substrato que eventualmente poderiam estar sendo transferidas para a variedade linguística em formação na situação de contato. (...)

Portanto, a maior diferença entre os processos típicos de pidginização e criouliização, que representam os casos mais radicais de transmissão linguística irregular, e os processos de transmissão linguística irregular mais leve, como o caso da formação das variedades populares da língua portuguesa no Brasil, é que, no primeiro caso, a gramática da variedade linguística que se forma na situação de contato é formada basicamente por elementos exógenos, enquanto, no segundo caso, os elementos gramaticais da língua do grupo dominante suplantam eventuais processos embrionários de gramaticalização e de transferência do substrato. (Lucchesi 2008: 371-372)

Enquanto as línguas pidgins e crioulas têm uma gramática qualitativamente distinta da gramática de sua língua lexificadora (Rougé 2008), as variedades formadas em situação de transmissão linguística irregular de tipo leve incorporariam a gramática da língua dominante na situação de contato. O que os dois processos teriam em comum seria a erosão morfológica que afetaria os mecanismos gramaticais sem valor informacional, como as regras de concordância e a flexão de caso. Nesse caso, entretanto, haveria uma diferença de grau, já que, no geral, esses elementos seriam completamente eliminados na pidginização/criouliização, ao passo que, na transmissão ligar irregular de tipo leve, o resultado seria um quadro de ampla variação no uso desses mecanismos gramaticais.

Análises tradicionais da primeira metade do século XX, como a de Chaves de Mello (1946), procuraram relacionar essa simplificação da morfologia verbal e nominal nas variedades do português popular à ausência desse mecanismo flexional nas línguas do substrato africano. Tal visão encontra respaldo em teorias mais recentes do substrato (cf. seção anterior). Porém, tal hipótese é contrariada pelos resultados alcançados por pesquisas muito bem fundamentadas empiricamente, levadas a cabo nas últimas décadas do século XX, sobre as variedades de segunda língua de trabalhadores imigrantes na Europa Ocidental e nos EUA. Essas análises têm demonstrado que a perda da morfologia flexional nas variedades de interlíngua que se formam quando um grande número de falantes adultos adquire uma segunda língua, de forma assistemática, com limitado acesso

aos modelos dessa língua alvo e sem a intenção de atingir uma plena proficiência nessa língua, ocorre independentemente das características estruturais da língua nativa desses falantes adultos, ou seja, dos modelos gramaticais do substrato, como postula Lefebvre, através da ideia de relexificação com *etiquetagem nula* (cf. seção anterior). Muitos desses trabalhadores eram falantes de línguas flexionais, como o espanhol e o turco, mas a variedade de interlíngua que desenvolviam, denominada *Variedade Básica*, carecia de elementos gramaticais mais abstratos, entre eles, os mecanismos morfossintáticos da concordância nominal e verbal (Klein e Perdue 1997).

Portanto, a erosão morfológica que caracteriza as variedades populares do PB pode ser atribuída aos processos que atuam na aquisição de uma segunda língua de forma defectiva por parte de uma grande população de falantes adultos, independentemente das características estruturais da gramática de suas línguas nativas (Lucchesi 2012b). Com efeito, os elementos gramaticais dos crioulos do Caribe e da Costa Ocidental da África que são mais consistentemente relacionados a processos de transferência do substrato africano, como a posposição dos determinantes, a serialização verbal, a formação de partículas pré-verbais de tempo modo e aspecto e a utilização do pronome da 3ª pessoa do plural como pluralizador nominal, não são encontradas no português popular brasileiro. A alternância dativa seria, em princípio, um candidato a possível influência estrutural do substrato. Porém, tal processo de variação e mudança poderia ser igualmente enquadrado no processo mais geral de simplificação morfológica que atua na gênese das línguas crioulas independentemente das influências do substrato, como já argumentado aqui, até porque essa é uma característica geral das línguas crioulas ao redor do mundo, mesmo daquelas línguas que não têm substrato africano (Lucchesi e Mello 2009). Dessa forma, a identificação de características estruturais do português popular brasileiro que possam ser associadas, de forma consistente, a uma possível influência de substrato é o novo desafio para aqueles que se colocam dentro da perspectiva que coloca o contato entre línguas como um fator crucial na formação histórica das variedades populares do português brasileiro.

4. CONCLUSÃO

No cenário de uma transmissão linguística irregular de tipo leve, como aqui referido, eventuais processos de transferência de substrato seriam particulares e laterais, ou seja, não seria provável encontrar transferências de substrato em estruturas e mecanismos nucleares da gramática das variedades populares do português brasileiro. Contudo, não devem ser descartadas investigações que busquem identificar remanescentes de processos pretéritos de transferência de substrato na história sociolinguística do Brasil. Nessa linha de investigação, há duas características de variedades populares do português no Brasil que parecem ser atualmente boas apostas, dentro da perspectiva da hipótese do substrato.

A primeira está ligada à estrutura argumental mais ampla do verbo *comprar*. Lumsden (1999a: 137-138) identifica uma construção do haitiano que tem um nítido paralelo com o fongbe, como se pode ver nos exemplos (24) e (25) abaixo, com o correspondente em português apresentado em (27). Ou seja, trata-se de um caso especial de regência verbal, em que o verbo *comprar* seleciona a locução prepositiva *na mão de* para reger seu argumento preposicionado. Essa construção não existe em francês, que marca esse argumento do verbo *acheter* com a preposição *de*, como ocorre no português padrão do Brasil – cf. exemplo (26) abaixo. Com base nisso, Lumsden apresenta esse fato como um caso de relexificação de uma palavra referencial do haitiano, que tem a forma fonética do francês, mas a especificação gramatical abstrata do substrato fongbe.

24 *Jan achte liv la nan men Pól.*

Jean comprar livro DEF em mão Paul

25 *Jan xô wémà ó dò Pòlù si.*

Jean comprar livro DEF em Paul mão

26 *Jean a acheté le livre de Paul.*

João comprou esse livro de Paulo.

27 João comprou esse livro **na mão** de Paulo.

Ocorre que a construção apresentada em (27) é muito comum na linguagem coloquial da Bahia, sendo desconhecida e mesmo agramatical em muitos outros estados do Brasil. Trata-se de um paralelo expressivo que merece ser investigado. O primeiro passo seria saber se a mesma construção existe no yoruba, a língua mais importante do substrato africano na Bahia. É bem provável que exista, já que as línguas são da mesma família linguística kwa. Além disso, o ewe e o fon, ambas do grupo fongbe, são línguas presentes no substrato africano, tanto na Bahia, quanto em Minas Gerais, chegando a predominar no Maranhão. Se a construção não existir no yoruba, pode-se recorrer ao esquema do *nivelamento dialetal*, como proposto por Lefebvre (cf. seção 1 supra), para explicar sua transferência para o português popular da Bahia. Outra evidência crucial seria fornecida pela distribuição do fenômeno no território brasileiro, o que poderia ser feito por meio de uma investigação nos materiais do ALIB (cf. seção 2 supra). O cenário ideal seria se essa construção estivesse restrita aos Estados de Minas Gerais, Bahia e Maranhão, para os quais se tem registros históricos de forte presença de substrato kwa, ou mais especificamente fongbe. Mas não se podem descartar eventuais processos históricos de difusão dessa construção para outros estados brasileiros, o que demandaria uma pesquisa histórica suplementar para atestar a ocorrência de tal difusão.

O segundo fato que sugere uma transferência do substrato também advém de um paralelo com o que foi observado no haitiano por Lumsden (1999a: 138-9). Trata-se de construções como: *pye-bananann* (lit. *pé de banana*), *pye-palmis* (lit. *pé de palma*) e *pye-bwa* (lit. *pé de árvore*). Essas construções não encontram paralelo na língua lexificadora, o francês, que expressa esses conteúdos pelo mecanismo da derivação (como ocorre no português padrão): *ba-*

nanier (*bananeira*), *palmier* (*palmeira*); sendo que *arbre* (*árvore*) não aceita tal derivação. Em contrapartida, construções similares são encontradas em fongbe: *kwékwé-tín* (lit. *árvore de banana*) e *dè-tín* (lit. *árvore de palma*). Não ocorre, entretanto, em fongbe um correspondente exato, tomando por base o substantivo *pé*.

Mais uma vez existe um paralelo notável com construções típicas da linguagem coloquial da Bahia: *pé de banana* por *bananeira*;⁷ *pé de palma* por *palmeira*, e mesmo *pé de árvore*, ou *pé de pau*. E, da mesma forma, tais construções inexistem em outros estados brasileiros. O caminho adotado por uma investigação empírica para verificar essa provável origem no substrato africano seria o mesmo proposto para a regência especial do verbo *comprar*: (i) verificar a existência de construções paralelas em yoruba e (ii) verificar a distribuição diatópica do fenômeno com base nos materiais do ALIB.

Esses dois possíveis candidatos a influências de substrato enquadram-se na visão geral de formação histórica do português popular brasileiro aqui assumida, na qual predomina o processo de *transmissão linguística irregular de tipo leve*, pois se trata de alterações em itens lexicais, e não em itens funcionais, que se situam no núcleo da gramática. Nas situações de transmissão linguística irregular de tipo leve, as influências do substrato seriam, como se disse acima, particulares e laterais.

De qualquer forma, a questão da influência dos substratos na história sociolinguística do Brasil ainda vai demandar muita reflexão teórica e pesquisa empírica. Trata-se de mais uma profícua frente de investigação que está se abrindo para a pesquisa linguística no Brasil, na medida em que vão ficando para trás as tradicionais visões imanentistas que buscavam restringir a formação histórica das variedades do português brasileiro à ação de tendências seculares prefiguradas na estrutura interna da língua.

7 Para além dessa forma, a variante mais corrente no português brasileiro é *pé de bananeira*, e construções afins, como *pé de laranjeira* etc.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- Arends, Jacques. 1994. The socio-historical background of creoles, em J. Arends, P. Muysken e N. Smith (eds.), *Pidgins and creoles: an introduction*, Amsterdam/Philadelphia, John Benjamins: 15-24.
- Baxter, Alan. 1988. *A Grammar of Kristang (Malacca Creole Portuguese)*, Canberra, Pacific Linguistics.
- Baxter, Alan e Dante Lucchesi. 1997. A relevância dos processos de pidginização e criolização na formação da língua portuguesa no Brasil, *Estudos Linguísticos e Literários*, 19: 65-84.
- Bickerton, Derek. 1981. *Roots of language*, Ann Arbor, Karoma.
- Bickerton, Derek. 1999. How to acquire language without positive evidence: what acquisitionists can learn from Creoles? em M. Degraff (ed.), *Language creation and language change: creolization, diachrony, and development*, Cambridge, The MIT Press: 49-74.
- Cardoso, Suzana Alice e Jacyra Mota. 2012. Projeto Atlas Linguístico do Brasil: antecedentes e estágio atual, *Alfa: Revista de Linguística* (UNESP. Online), 56: 855-870.
- Castro, Yeda Pessoa de. 2002. *A língua mina-jeje no Brasil: um falar africano em Ouro Preto do século XVIII*, Belo Horizonte, Secretaria de Estado da Cultura.
- Chaves de Melo, Gladstone. 1946. *A língua do Brasil*, Rio de Janeiro, Agir.
- Chomsky, Noam. 1981. *Lectures on government and binding*, Dordrecht, Foris.
- Field, Fred. 2004. Second Language Acquisition in Creole Genesis, em G. Escure e A. Schwegler (eds.), *Creoles, Contact and Language Change: Linguistic and Social Implications*, Amsterdam/Philadelphia, John Benjamins: 127-160.
- Guy, Gregory. 1981. *Linguistic variation in Brazilian Portuguese: aspects of phonology, syntax and language history*, Tese de Doutorado, University of Pennsylvania, Pennsylvania. Inédita.
- Guy, Gregory. 1989. On the nature and origins of vernacular Brazilian Portuguese, em *Estudios sobre Español de América y Linguística Afroamericana*, Bogotá, Instituto Caro y Cuervo: 226-244.

- Goodman, Morris. 1964. *A Comparative Study of Creole French Dialects*, The Hague, Mouton.
- Holm, John. 1987. Creole influence on Popular Brazilian Portuguese, em G. G. Gilbert (ed.), *Pidgin and Creole languages. Essays in Memory of John E. Reinecke*, Honolulu, University of Hawaii Press: 406-429.
- Holm, John. 1992. Popular Brazilian Portuguese: a semi-creole, em E. D'Andrade e A. Kihm, (orgs.). *Actas do Colóquio sobre Crioulos de Base Lexical Portuguesa*, Lisboa, Colibri: 37-66.
- Keesing, Roger. 1988. *Melanesian Pidgin and the Oceanic Substrate*, Stanford, Stanford University Press.
- Klein, Wolfgang e Clive Perdue. 1997. The basic variety (or couldn't natural languages be much simpler?), *Second Language Research*, London, 13(4): 301-347.
- Koopman, Hilda. 1986. The genesis of Haitian: Implications of a comparison of some features of the syntax of Haitian, French and West African Languages, em P. Muysken e N. Smith (eds.), *Substrata Versus Universals in Creole Genesis*, Amsterdam/Philadelphia, John Benjamins: 231-258.
- Lefebvre, Claire. 1997. Relexification in Creole genesis: The case of demonstrative terms in Haitian Creole, *Journal of Pidgin and Creole Languages*, 12(2): 181-201.
- Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar: the case of Haitian Creole*, Cambridge, Cambridge University Press.
- Lefebvre, Claire. 2001. Relexification in creole genesis and its effects on the development of the creole, em N. Smith e T. Veenstra (eds.), *Creolization and Contact*, Amsterdam, John Benjamins: 9-42.
- Lefebvre, Claire e John Lumsden. 1992. On word order in relexification. *Travaux de recherche sur le créole haïtien*, Université du Québec à Montréal, n. 10, p. 1-22, 1992.
- Lucchesi, Dante. 2002. Grandes Territórios Desconhecidos, *Linguística*, 14: 191-222.
- Lucchesi, Dante. 2004. A participação do contato entre línguas na formação do português popular do Brasil: novas evidências empíricas, em M. Fernández, M. Fernández-Ferreiro e N. V. Veiga, *Los criollos de base ibérica*, Madrid, Iberoamericana: 215-226.

- Lucchesi, Dante. 2008. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto, em C. Roncarati e J. Abraçado (orgs.), *Português brasileiro II: contato lingüístico, heterogeneidade e história*, Niterói, EDUFF: 366-390.
- Lucchesi, Dante. 2009. História do Contato entre Línguas no Brasil, em D. Lucchesi, A. Baxter e I. Ribeiro (orgs.), *O Português Afro-Brasileiro*, Salvador, EDUFBA: 41-73.
- Lucchesi, Dante. 2012a. A diferenciação da língua portuguesa no Brasil e o contato entre línguas, *Estudos de Lingüística Galega*, Santiago de Compostela, 4: 45-65.
- Lucchesi, Dante. 2012b. *A deriva secular na formação do português brasileiro: uma visão crítica*, em T. Lobo, Z. Carneiro, J. Soledade, A. Almeida, S. Ribeiro (orgs.), *ROSAE: lingüística histórica, história das línguas e outras histórias*, Salvador, EDUFBA: 249-274.
- Lucchesi, Dante. 2013. O contato entre línguas e a origem do português brasileiro, em E. Gugenberger, H. Monteagudo e G. Rei-Doval, *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias*, Santiago de Compostela, Consello da Cultura Galega: 191-221.
- Lucchesi, Dante; Alan Baxter e Ilza Ribeiro (orgs.). 2009. *O Português Afro-Brasileiro*, Salvador, EDUFBA.
- Lucchesi, Dante e Camila Mello. 2009. A alternância dativa, em D. Lucchesi, A. Baxter e I. Ribeiro (orgs.), *O Português Afro-Brasileiro*, Salvador, EDUFBA: 427-456.
- Lumsden, John. 1999a. Language acquisition and creolization, em M. DeGraff (ed.), *Language creation and language change: creolization, diachrony, and development*, Cambridge, The MIT Press: 129-157.
- Lumsden, John. 1999b. The role of relexification in creole genesis, *Journal of Pidgin and Creole Languages*, Amsterdam, 14(2): 225-258.
- Mattoso, Katia. 2003. *Ser escravo no Brasil*, 3ª ed., São Paulo, Brasilien-se.
- McWhorter, John. 1998. Identifying the creole prototype. Vindicating a typological class, *Language*, 74 (4): 788-818.
- McWhorter, John. 2001. The world's simplest grammars are creole grammars, *Linguistic Typology*, 5 (2/3): 125-166.
- Mendonça, Renato. 1933. *A influência africana no português do Brasil*, Rio de Janeiro, Sauer.

- Mühlhäusler, Peter. 1986. *Pidgin & Creole linguistics*, Londres, Basil Blackwell.
- Queiroz, Sônia. 1998. *Pé preto no barro branco: a língua dos negros da Tabatinga*, Belo Horizonte, EDUFMG.
- Rampton, Ben. 1997. A sociolinguistic perspective on L2 communication strategies, em G. Kasper e E. Kellerman (eds.), *Communication Strategies: Psycholinguistic and Sociolinguistic Perspectives*, Londres/Nova York, Longman: 279-303.
- Reesink, Ger. 1990. Mother tongu and Tok Pisin, em J. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin*, Amsterdam, John Benajmins: 289-306.
- Ribeiro, Darcy. 1995. *O povo brasileiro: evolução e sentido do Brasil*, São Paulo, Companhia das Letras.
- Roberts, Sarah. 2000. Nativization and genesis of Hawaiian Creole, em J. McWhorter (ed.), *Language Change and Language Contact in Pidgins and Creoles*, Amsterdam/Philadelphia, Benjamins: 333-352.
- Rodrigues, Nina. [1993] 2004. *Os africanos no Brasil*, 8ª Ed, Brasília, Editora Universidade de Brasília,
- Rougé, Jean Louis. 2008. A inexistência de crioulo no Brasil, em J. L. Fiorin e M. Petter (eds.), *África no Brasil: a formação da língua portuguesa*, São Paulo, Contexto: 63-74.
- Shumann, John. 1978. *The Pidginization Process: a model for a second language acquisition*, Rowley: Newbury House,
- Siegel, Jeff. 2008. *The Emergence of Pidgin and Creole Languages*, Oxford, Oxford University Press.
- Sylvain, Suzanne. 1936. *Le créole haïtien: Morphologie et syntaxe*, Port-au-Prince, edição do autor.
- Voorhoeve, Jan. 1973. Historical and linguistic evidence in favor of the relexification theory in the formation of creoles, *Language and Society*, 2: 133-145.
- Wekker, Herman. 1996. Creolization and the acquisition of English as a second language, em W. Herman (ed.), *Creole Languages and Language Acquisition*, Berlin, Mouton de Gruyter: 139-149.
- Whinnom, Keith. 1971. Linguistic hybridization and the “special case” of pidgins and creoles, em D. Hymes (ed.), *Pidginization and creolization of languages*, Cambridge, Cambridge University Press: 91-115.