

A SIMPLIFICAÇÃO MORFOLÓGICA NA EXPRESSÃO DO SUJEITO INDETERMINADO NO PORTUGUÊS AFRO-BRASILEIRO

por Dante Lucchesi (UFBA)¹

RESUMO

Uma frente da pesquisa linguística no Brasil busca os efeitos do contato entre línguas na formação do português brasileiro. Um desses efeitos é a utilização da forma verbal da terceira pessoa do singular para expressar a indeterminação do sujeito. Esse processo de simplificação morfológica enquadra-se no tipo de mudança que ocorre em determinadas situações de contato linguístico massivo, presentes na história sociolinguística do Brasil. Para testar essa hipótese, são analisadas amostras sociolinguísticas de fala vernácula de comunidades rurais afro-brasileiras do interior do Estado da Bahia.

PALAVRAS-CHAVE: sujeito indeterminado; contato entre línguas; variação linguística; português afro-brasileiro.

MORPHOLOGICAL SIMPLIFICATION IN THE REPRESENTATION OF INDEFINITE SUBJECTS IN AFRO-BRAZILIAN PORTUGUESE

ABSTRACT

One line of linguistic research in Brazil seeks of effects of language contact in the development of Brazilian Portuguese. One of these effects is the use of the third person singular form of the verb to express the indeterminacy of the subject. This morphological simplification fits the type of change that occurs in certain situations of massive language contact, such as occurred in the sociolinguistic history of Brazil. To test this hypothesis, this work analyzes sociolinguistic samples of vernacular speech of African-Brazilian rural communities in the state of Bahia.

KEY WORDS: undefined subject; language contact; linguistic variation; Afro-Brazilian Portuguese.

1. Professor Titular de Língua Portuguesa da Universidade Federal da Bahia, Bolsista de Produtividade em Pesquisa do CNPq, Doutor em Linguística pela Universidade Federal do Rio de Janeiro e Mestre em Linguística Portuguesa Histórica pela Universidade de Lisboa.

1. INTRODUÇÃO

O avanço da visão que destaca a proeminência do contato entre línguas na formação histórica do português brasileiro tem estimulado a pesquisa sobre a extensão dos efeitos desse processo histórico na estrutura gramatical das variedades linguísticas brasileiras que teriam sido mais diretamente afetadas pelo contato do português com as línguas indígenas e africanas.² Um desses potenciais efeitos é a utilização da forma verbal da 3ª pessoa do singular com sujeito nulo e sem referência anterior, para expressar a indeterminação do sujeito, como se pode ver no seguinte exemplo:

(1) Hoje **usa** mais roupa colorida do que antigamente.

Essa construção pode ser vista como o resultado de uma simplificação morfológica em que se suprime a partícula gramatical *se*, presente na construção canônica:

(2) Hoje **se usa** mais roupa colorida do que antigamente.

Considerando que a simplificação morfológica é a principal consequência das situações de contato massivo, que têm na *pidginização* e na *crioulização* suas manifestações mais radicais (SIEGEL, 2008), a predominância dessa construção nas variedades populares do português brasileiro poderia ser vista como um reflexo de mudanças induzidas pelo contato entre línguas que teriam ocorrido na formação histórica dessas variedades linguísticas, quando o português foi adquirido como segunda língua, em situações adversas, por milhões de índios aculturados e africanos escravizados, tornando-se progressivamente a língua materna de seus descendentes; o que configura o que temos denominado *transmissão linguística irregular de tipo leve*, em oposição à transmissão linguística irregular mais radical que ocorreria nos processos de pidginização e crioulização típicos (LUCCHESI, 2003 e 2008a).

Para testar essa hipótese, apresentamos aqui os resultados de uma análise sociolinguística sobre esse fenômeno na fala de quatro comunidades rurais afro-brasileiras do interior do Estado da Bahia (algumas delas oriundas de antigos quilombos), baseada nos pressupostos da Teoria da Variação Linguística (LABOV, 1972 e 1994). Mas, além da apresentação dos resultados dessa análise, este artigo contém: uma primeira seção, que trata do tema do contato entre línguas na história sociolinguística do Brasil; uma segunda seção, que se debruça sobre o fenômeno linguístico da indeterminação do sujeito nas línguas em geral e nas línguas crioulas em particular; e uma terceira seção, que descreve a variação na expressão gramatical do sujeito indeterminado no português brasileiro.

2. Agradeço aos pareceristas anônimos da Revista pelos comentários e sugestões, que foram incorporadas, na medida do possível, nesta versão final do artigo. Entretanto, eventuais lapsos e incorreções são de minha inteira responsabilidade.

2. O CONTATO ENTRE LÍNGUAS NA HISTÓRIA SOCIOLINGUÍSTICA DO BRASIL: A TRANSMISSÃO LINGUÍSTICA IRREGULAR DE TIPO LEVE

A colonização do Brasil baseou-se desde sempre no extenso emprego do trabalho forçado. Inicialmente, os colonizadores portugueses recorreram ao apresamento dos indígenas. Porém, já no século XVI, iniciou-se a importação de escravos africanos. Assim, por quase quatro séculos, se opuseram, na composição da sociedade brasileira, uma minoria branca, de falantes nativos da língua portuguesa, constituindo o segmento dominante, e uma larga maioria de índios aculturados e africanos escravizados, falantes de centenas de línguas diversas e ininteligíveis entre si, compondo a grande massa de explorados. Em função da dominação física e simbólica, os descendentes de índios e africanos abandonavam a língua de seus ascendentes e adotavam a língua do colonizador como língua materna. Portanto, o cenário sociolinguístico do período que se estende do início da colonização até o início da república define-se por uma *diglossia* entre a língua dos colonizadores (fortemente influenciada pelos modelos linguísticos da metrópole portuguesa) e as línguas africanas e indígenas, sucedidas por variedades bastante alteradas do português, faladas por africanos e indígenas, como segunda língua, ou como língua materna, por seus descendentes (LUCCHESI, 2009a). Radicam nessa diglossia as bases históricas da atual *polarização sociolinguística do Brasil*, que opõe, em seus extremos, a norma linguística da elite letrada à fala de uma grande massa de brasileiros que ainda hoje são economicamente marginalizados e excluídos do espaço da cidadania (LUCCHESI, 2001, 2002 e 2006).

Embora o contexto sócio-histórico do Brasil colonial, nomeadamente no século XVII, tenha sido favorável à emergência de línguas pidgins e crioulas (GUY, 1981), os processos de pidginização e crioulição do português teriam ocorrido de forma efêmera e localizada (LUCCHESI, 2009a). Apesar disso, não é razoável pensar que o contato entre línguas tenha desempenhado um papel secundário na formação das variedades populares do português brasileiro contemporâneo. Até o final do século XIX, cerca de dois terços da população do Brasil era composta por índios e africanos e seus descendentes, que adquiriam a língua portuguesa nas situações as mais adversas; sendo que cerca de noventa por cento da população vivia fora dos centros urbanos e um percentual ainda maior vivia fora do universo do letramento. Nesse contexto sócio-histórico, a formação das atuais variedades populares do português brasileiro ocorreu a partir da aquisição precária do português por parte de milhões de índios aculturados e africanos escravizados, em sua larga maioria adultos, seguida da nativização desse modelo defectivo de segunda língua entre os seus descendentes. Embora os falantes do antecedente histórico do português popular contemporâneo tenham tido mais acesso aos modelos gramaticais do português nativo do que o que acontece nas situações típicas de crioulição, em função de características específicas da sociedade brasileira, o multilinguismo maciço e o incipiente letramento criaram as condições favoráveis para alterações estruturais significativas nas variedades populares do português brasileiro (MATTOS E SILVA, 2004).

Tomando por base os avanços, nas últimas décadas, das teorias sobre aquisição de segunda língua e sobre as situações de contato linguístico massivo e radical (WEKKER, 1996; FIELD, 1997; KLEIN; PERDUE, 1997; SIEGEL, 2008), temos argumentado no sentido de que as principais características que contrastam atualmente a norma popular brasileira com a norma urbana culta (como o amplo processo

de variação no uso das regras de concordância nominal e verbal) encontram suas origens em processos de mudança induzidos pelo contato entre línguas, no que temos denominado processo de *transmissão linguística irregular de tipo leve* (LUCCHESI, 2012 e 2013; LUCCHESI; BAXTER, 2006).

O conceito de *transmissão linguística irregular* é um conceito amplo e gradual que visa a dar conta de uma gama de situações de contato entre línguas, recobrando, tanto as situações típicas de *crioulização*, que dão ensejo à formação de variedades linguísticas qualitativamente distintas da língua dominante na situação de contato (ROUGÉ, 2008), quanto as situações menos radicais, como as que estão na origem das variedades populares do português do Brasil, que sofreram mudanças estruturais em função do contato entre línguas, mas não com um grau de radicalidade suficiente para formar uma gramática qualitativamente distinta da gramática do português.

No geral, crioulização implica a eliminação de certos mecanismos gramaticais sem valor referencial, como as regras de concordância nominal e verbal, o movimento e a flexão de caso, ao lado da recomposição de outros mecanismos gramaticais da estrutura nuclear da gramática, como o sistema verbal de tempo, modo e aspecto, os conectivos oracionais e sintagmáticos e as partículas gramaticais para indicar reflexividade e indeterminação do sujeito. Essa recomposição gramatical tem por base a transferência estrutural das línguas dos grupos dominados (LEFEBVRE, 1998; LUMSDEN, 1999; SIEGEL, 2008) ou os dispositivos da faculdade humana da linguagem que atuam na aquisição da língua materna (BICKERTON, 1981 e 1999). Tanto num caso quanto noutro, é essa reestruturação gramatical que garante à língua crioula uma gramática qualitativamente distinta da língua da qual herdou praticamente todo o seu léxico, a língua do grupo socialmente dominante, denominada, por essa razão, *língua lexicadora* ou língua de superestrato.

A crioulização típica ocorre em situações históricas muito específicas, em que povos substituem sua língua nativa por outra, à qual têm muito pouco acesso, tanto que a grande maioria das línguas crioulas é falada por populações que são ou foram muito marginalizadas. Porém, em situações históricas, como a colonização do continente americano a partir do século XVI, ocorreram situações de contato massivo entre línguas, em contextos sociais muito mais amplos do que os contextos típicos de crioulização. Nessas situações, os falantes das outras línguas e, sobretudo, seus descendentes tiveram mais acesso aos modelos gramaticais da língua dominante, inibindo, assim, potenciais processos de transferência de substrato e de gramaticalização, que são cruciais para a formação das línguas crioulas.

Porém, deve-se ter em conta que a erosão gramatical da língua dominante não deixa de ocorrer mesmo nesse contexto mais amplo. Essa erosão mais leve atinge, então, os mecanismos gramaticais sem valor referencial, como, por exemplo, as regras de movimento e de concordância nominal e verbal, não a ponto de eliminá-las, mas produzindo um amplo quadro de variação em seu emprego. Além disso, há outra diferença crucial entre a transmissão linguística irregular de tipo leve e a crioulização: os processos de gramaticalização e reestruturação original da gramática, típicos desta, são raros ou ausentes naquela.

As análises que temos conduzido sobre os processos de variação que diferenciam as variedades populares das variedades cultas do português brasileiro obedecem ao previsto nessa visão. No caso

específico da variação na expressão da indeterminação do sujeito, a expectativa não é encontrar um mecanismo gramatical original de expressão dessa categoria gramatical, quiçá proveniente da gramática de uma língua do substrato africano, como ocorre, por exemplo, em alguns crioulos de base lexical portuguesa da África. O mais provável é que se configure um processo de simplificação morfológica, que pode ter um paralelo em crioulos mais leves. Assim, o ponto de partida é sempre a observação de como esse mecanismo gramatical é afetado nos processos típicos de crioulição.

3. A INDETERMINAÇÃO DO SUJEITO NAS LÍNGUAS HUMANAS EM GERAL E NAS LÍNGUAS CRIOULAS EM PARTICULAR

Pode-se assumir que a programação gramatical das línguas humanas contém potencialmente um dispositivo para expressar a *indeterminação do sujeito*. Nas línguas românicas, o pronome reflexivo/recíproco *se* funciona também como índice de indeterminação do sujeito:

- (3) a. Aqui *se* vive bem. (português)
- b. *Qui se vive bene*. (italiano)
- c. *Se vive bien aquí*. (espanhol)

Isso não significa que todas as línguas devam exibir esse dispositivo em seu repertório gramatical, já que esse, como qualquer outro conteúdo gramatical, pode ser expresso lexicalmente:

- (4) *Qualquer pessoa vive bem aqui*.

Porém, mesmo no plano da gramática, há formas alternativas de expressar o sujeito de referência genérica, como a 1ª pessoa do plural, expressa através do pronome ou apenas pela desinência verbal, como nos exemplos abaixo:

- (5) a. Devemos respeitar os mais velhos.
- b. *Dobbiamo rispettare i piú vecchi*.
- c. *Debemos respetar a los ancianos*.

Já no inglês, a forma mais corrente de expressar a indeterminação do sujeito se dá com o recurso ao numeral *one* – cf. exemplo (6) –, que se gramaticalizou nessa função. Mas os pronomes de 2ª pessoa do singular e da 1ª pessoa do plural também podem assumir essa função, como ocorre em português – cf. exemplos (7) e (8), respectivamente:

- (6) *One lives well here*.
- (7) *You should respect the elderly*.
- (8) *We should respect the elderly*.

Em alemão, a gramaticalização ocorreu com o substantivo *mann* ‘homem’ que assume a forma *man* nessa função – cf. exemplos (9) e (10).³ Em francês, observa-se um fenômeno similar com a forma

3. Agradeço a Iris Rieder pelas informações sobre o alemão.

on – cf. exemplo (11) –, que tem sua origem no substantivo **homo** do latim, que era uma das formas de expressar o sujeito de referência genérica, no chamado *latim vulgar*.⁴

(9) *Man lebt hier gut.*

homem vive aqui bem

Vive-se bem aqui.

(10) *Man muss die Älteren respektieren.*

homem deve os velhos respeitar.

Devemos respeitar os mais velhos.

(11) *On doit respecter les plus vieux.*

Devemos respeitar os mais velhos.

Já no chamado latim clássico ou literário, a passiva seria a forma mais usual de expressar a indeterminação do sujeito, como se pode ver na frase (12) abaixo, adaptada de uma fala de uma peça de Plauto. Porém a forma verbal da 1ª pessoa do plural também era usada nessa função, como no exemplo (13), extraído de um texto do filósofo Sêneca:⁵

(12) *Hic bene vivitur.*

aqui bem viver-Pass

Vive-se bem aqui.

(13) *Apes debemus imitari.*

abelha-Ac-Pl devemos imitar

Devemos imitar as abelhas.

No que diz respeito às formas gramaticais de expressar a indeterminação do sujeito nas línguas crioulas, não há muita informação disponível na literatura do campo. Bickerton (1981), por exemplo, não inclui a forma gramatical de expressão do sujeito de referência genérica entre os doze traços virtualmente presentes nas línguas crioulas ao redor do mundo. Há muitas referências à gramaticalização dos substantivos *cabeça* e *corpo* como pronome reflexivo, como ocorre, por exemplo, no caboverdiano e no santomense, respectivamente, mas não há registro da gramaticalização do substantivo *homem*, por exemplo, como índice de indeterminação do sujeito, em línguas crioulas.

Há, entretanto, o registro feito por Luiz Ivens Ferraz (1979, p. 66) de um “pronome pessoal não marcado para pessoa e número” *a* no forro ou santomense, crioulo de base lexical portuguesa, que é a língua materna da maioria da população da Ilha de São Tomé, que forma a república insular de São

4. A gramaticalização do substantivo *homem* como forma para expressar a indeterminação do sujeito também ocorreu, na fase arcaica do português e de várias outras línguas românicas, como o espanhol, o provençal e o italiano, mas só se conservou no francês. Em português, a lacuna deixada pelo desuso do substantivo *homem* (*ome*), com essa função, foi preenchida com o emprego da forma nominal *a gente* (LOPES, 2003).

5. Agradeço a José Amarante pelas informações sobre o latim.

Tomé e Príncipe, no Golfo da Guiné, na costa ocidental do continente africano. Segundo Ferraz, o *a* equivale ao pronome *se* do português, na função de índice de indeterminação do sujeito, como se pode ver no exemplo abaixo:

(14) *A ka pó fé kwa sé.*

Indef Pres pode fazer coisa Def
'Pode-se fazer isso.'

Segundo Ferraz, trata-se de uma transferência desse pronome da língua edô, da família linguística kwa e falada no Benin, que é considerada a principal língua do substrato dos crioulos portugueses do Golfo da Guiné. O fenômeno está presente nos outros três crioulos portugueses do Golfo da Guiné, que têm uma origem comum com o santomense: o principense ou Lung'ie, falado na Ilha do Príncipe; o angolar, falado ao sul da Ilha de São Tomé; e o annabonense ou fa d'Ambô, falado na ilha de Annobón, (HAGEMEIJER, 2009, p. 14); como se pode ver neste exemplo do principense:

(15) *A sa kunfya na mye fa*

Indef Neg confiar em mulher Neg
'Não se deve confiar nas mulheres.'

Já no papiá kristang, crioulo português falado em Malaca, na Malásia, país do sudeste asiático, a indeterminação do sujeito é expressa através da gramaticalização do substantivo *gente*:

(16) *Jénti m'pódi komfiá kù mulé.*

Indef Neg-pode confiar Ac mulher
'Não se pode confiar nas mulheres.'

Porém, em alguns contextos, o sujeito nulo, sem referência anterior, é a construção utilizada para expressar a indeterminação do sujeito, como se pode ver nos exemplos abaixo:⁶

(17) *Miste respétu ku jénti idádi.*

deve respeito Ac gente idade
'Devemos respeitar os mais velhos.'

(18) *Témpu di samatra, m'pódi bai mar.*

tempo de tempestade, Neg-pode ir mar
'Com tempestade, não se pode ir para o mar.'

Apesar de não haver registros do emprego dessa estratégia de indeterminação do sujeito em outras línguas crioulas, é provável que ela seja mais geral, pois se combina, por um lado, com simplificação morfológica que caracteriza essas línguas (MCWHORTER, 1998, 2001; SIEGEL, 2008) e, por outro lado, com o fato de, na maioria dos crioulos, não ser possível o sujeito nulo com referência definida.

6. Agradeço a Alan Baxter pelas informações sobre o papiá kristang.

Portanto, nas línguas crioulas, ocorrem, em termos mais amplos, os mesmos mecanismos gramaticais de indeterminação do sujeito encontrados nas línguas não crioulas: tanto partículas gramaticais de significado menos transparente, como o *a* dos crioulos do Golfo da Guiné, que corresponderia ao *se* das línguas românicas; quanto formas que revelam um processo de gramaticalização ainda visível, como o *Jénti*, do crioulo português de Malaca, que corresponderia, grosso modo, ao *man*, do alemão, e ao *one*, do inglês. Mas haveria também a construção morfologicamente simplificada, em que o sujeito nulo expressa o sujeito de referência genérica, como encontrado no crioulo de Malaca. Tal construção, que deve ser mais geral entre as línguas crioulas, é crucial para a compreensão do quadro de variação observado hoje no português popular brasileiro, no que concerne ao papel desempenhado pelo contato entre línguas em sua formação.

4. A INDETERMINAÇÃO DO SUJEITO NO PORTUGUÊS BRASILEIRO EM GERAL E NO PORTUGUÊS POPULAR BRASILEIRO EM PARTICULAR

A tradição gramatical define o sujeito indeterminado como a situação em que “o verbo não se refere a uma pessoa determinada, ou por se desconhecer quem executa a ação, ou por não haver interesse no seu conhecimento”. A essa definição de cunho semântico agrega-se uma definição formal, em que se apresentam as duas formas de expressão do sujeito indeterminado (CUNHA, 1976, p. 141):

i) verbo na 3ª pessoa do plural

(19) **Roubaram** o celular da Maria.

ii) verbo na 3ª pessoa do singular, com o pronome *se*

(20) **Vive-se** bem aqui.

No segundo caso, restringe-se essa possibilidade aos verbos intransitivos ou que “tragam complemento preposicional” (ROCHA LIMA, 1960, p. 226-7); isso porque a construção com verbos transitivos diretos é interpretada como *voz passiva sintética*, pois essa construção sintética, exemplificada em (21), equivaleria a construção analítica, exemplificada em (22):

(21) **Avistou-se** um monte na época da Páscoa.

(22) Um monte **foi avistado** na época da Páscoa.

Porém, diversas pesquisas, como as de Nunes (1990) e Cavalcante (1999), têm revelado que, no português brasileiro, o valor nominativo do *se* estendeu para as construções com os verbos transitivos diretos, de modo que todas as construções com o *se* poderiam ser interpretadas como indeterminação do sujeito. Com efeito, a falta de naturalidade da chamada *passiva sintética* é particularmente crítica nas construções com um auxiliar modal, como exemplificado em (23), em que a concordância com o constituinte nominal que sucede o verbo pode ser sentida como agramatical, mesmo por um brasileiro escolarizado.

(23) Devem-se respeitar os mais velhos.

Portanto, o problema na análise da tradição gramatical é a contradição na definição de sujeito indeterminado, entre o plano semântico e o formal. No plano do significado, a referência genérica ao agente da ação expressa pelo verbo incluiria a chamada *passiva sintética*, bem como as estruturas passivas em que esse agente não é expresso. Há, igualmente, nesse plano semântico, uma confusão entre diferentes níveis de referencialidade do sujeito que são agrupadas sob o mesmo rótulo de *sujeito indeterminado*:

i) sujeito de referência genérica universal

(24) Não **se** deve comer muito à noite.

ii) sujeito de referência parcialmente genérica

(25) Planta-**se** muito feijão nesta localidade.

(26) Nunca **ajudaram** o João.

iii) sujeito de referência específica indefinida

(27) Fechou-**se** o carro com chave dentro.

(28) Me **falaram** mal de você.

Há uma clara oposição entre os níveis (i) e (ii), de um lado, e o nível 3, de outro. A referência genérica, ou não específica, só cabe nos casos (i) e (ii), quando o sujeito se refere a todos os indivíduos de uma espécie, como em (24), ou a uma coletividade, como em (25), ou a um grupo de pessoas com um contorno impreciso, como em (26). Situação bem distinta é a que se verifica em (27) e (28), quando a referência é feita a um conjunto específico de indivíduos, que pode ser mesmo um conjunto unitário, um indivíduo, cuja identidade é irrelevante, ou desconhecida. Há uma diferença formal que reflete a distinção entre esses dois níveis de referencialidade. A forma da 3ª pessoa do plural não pode ser usada nos casos de referência genérica universal, sendo mais apropriada aos casos de sujeito específico indefinido. Já o emprego do *se* seria mais apropriado para os casos de referência genérica, embora também possa ser usado nos casos de sujeito específico indefinido, cf. exemplo (27).

Pode-se concluir, então, que o conceito de indeterminação do sujeito, na tradição gramatical, é essencialmente formal, circunscrevendo-se às duas formas arroladas como tal. Nesse caso, a tradição peca mais uma vez por ser excessivamente restritiva, ou seletiva, ignorando outras formas gramaticais empregadas no uso da língua para expressar a indeterminação do sujeito. Tal é o caso dos pronomes pessoais da 2ª pessoa do singular e da 1ª pessoa do plural, que podem assumir uma referência genérica, ou parcialmente genérica:

(29) **Você** sempre tem a sensação de esquecer alguma coisa, quando sai em viagem.

(30) **Devemos** respeitar os mais velhos.

(31) **A gente** estuda pouco gramática na Faculdade de Letras.

O pronome da 3ª pessoa do plural também pode ser empregado com referência parcialmente genérica ou com referência indefinida:

(32) **Eles** plantam muito feijão nesta localidade.

(33) O roubo foi muito bem planejado, **eles** sabiam o que estavam fazendo.

Na linguagem coloquial ou popular, emprega-se também a forma da 3ª pessoa do singular, sem qualquer pronome e fora de qualquer cadeia de correferência, para expressar o sujeito indeterminado:

(34) Hoje em dia, **usa** muito roupa colorida.

Essa construção é bastante significativa, se pensarmos em termos de um possível efeito do contato entre línguas na formação do português brasileiro, em função de um nítido paralelo que pode ser traçado com pelo menos uma língua crioula de base lexical portuguesa, que também emprega essa estratégia de indeterminação do sujeito (cf. seção anterior), divergindo do que ocorre no português de Portugal, que, em princípio, não aceitaria como gramatical tal construção. Dessa forma, essa construção pode ser vista como resultante de um processo histórico de simplificação morfológica, no qual se eliminou o emprego do pronome *se* como índice de indeterminação do sujeito, de modo que o sujeito nulo sem referência anterior, por si só, já seria suficiente para expressar o sujeito de referência genérica.

Vista dessa forma, como um processo de simplificação morfológica, a emergência da construção de sujeito nulo com a forma verbal da 3ª pessoa do singular sem o *se*, e sem referência anterior, como forma de expressão do sujeito indeterminado se enquadra, em linhas gerais, nos processos de mudança que podem ter sido induzidos pelo contato entre línguas nas situações de transmissão linguística irregular de tipo leve que caracterizam a história sociolinguística do Brasil, na forma como argumentamos na seção 1 deste artigo.

A próxima seção apresenta os resultados de uma análise sociolinguística da variação no emprego das formas gramaticais de expressão da indeterminação do sujeito na fala de quatro comunidades rurais afro-brasileiras do interior do Estado da Bahia. Essas comunidades seriam representativas da variedade linguística que supostamente seria a que mais foi afetada pelo contato entre português e as línguas africanas no Brasil.

5. ANÁLISE SOCIOLINGUÍSTICA DA VARIAÇÃO NAS FORMAS DE INDETERMINAÇÃO DO SUJEITO NO PORTUGUÊS AFRO-BRASILEIRO

A análise que aqui se apresenta da variação nas formas gramaticais de indeterminação do sujeito no que temos denominado *português afro-brasileiro* segue os pressupostos da Teoria da Variação Linguística, dentro da abordagem da mudança em *tempo aparente* (LABOV, 1972 e 1994). Para tanto, tomou-se como base empírica amostras de fala vernácula de quatro comunidades rurais afro-brasileiras isoladas de diferentes regiões do Estado da Bahia. Essas comunidades formam um dos objetos de investigação de um amplo programa de pesquisa que conduzimos na Universidade Federal da Bahia, denominado Projeto *Vertentes do Português Popular do Estado da Bahia*, ou apenas Projeto *Vertentes*.⁷

7. Mais informações sobre o Projeto *Vertentes* podem ser encontradas na página do projeto na Internet: <http://www.vertentes.ufba.br/>.

A primeira das quatro comunidades estudadas é geminada em duas, Barra e Bananal, conhecidas como *Arraiais de Rio de Contas*, município da Chapada Diamantina, região marcada em sua história pelo garimpo do ouro e pedras preciosas. Essas duas comunidades se originaram provavelmente em agrupamentos de escravos foragidos (quilombos), no século XVIII. A segunda comunidade, chamada *Cinzento*, situa-se no Município de Planalto, no semiárido baiano, onde predomina a pecuária. Também deve ter sido, em sua origem, um quilombo. A terceira comunidade, *Helvécia*, está situada no Município de Nova Viçosa, próximo ao litoral do extremo sul do Estado da Bahia, e teve a sua origem em uma colônia de alemães, suíços e franceses que, entre o início e o final do século XIX, estabeleceram na região grandes plantações de café que empregavam largamente a mão de obra de escravos africanos. Após a abolição da escravatura, os colonos europeus abandonaram o país deixando na terra os seus ex-escravos. A quarta comunidade, denominada *Sapé*, situa-se no Município de Valença, um pouco ao sul do chamado Recôncavo Baiano, cuja história está ligada aos grandes engenhos de cana-de-açúcar, que prosperaram entre os séculos XVII e XVIII. A comunidade foi formada também por ex-escravos que, nos finais do século XIX, se fixaram em terras mais ermas e recônditas que receberam de seus ex-senhores. É a comunidade que está mais próxima de Salvador, capital do Estado e grande centro de irradiação linguística.⁸

Em cada uma das quatro comunidades, foram gravadas entrevistas de tipo sociolinguístico com doze de seus membros, escolhidos aleatoriamente e distribuídos equitativamente entre os dois sexos e três faixas etárias: Faixa I, de 21 a 40 anos; Faixa II, de 41 a 60 anos, e Faixa III, mais de 60 anos. Os moradores entrevistados, assim como a maioria comunidade, tinham um nível de escolarização nulo ou quase nulo. A fala dessas comunidades assemelha-se ao português popular rural, podendo apresentar traços linguísticos ainda mais desviantes do que este, em relação ao português urbano culto. A análise quantitativa das ocorrências do fenômeno variável nessa amostra de fala utilizou o pacote de programas VARBRUL, em sua versão *GoldVarb X* (GUY; ZILLES, 2007).

Com esse enquadramento teórico-metodológico, a análise formalizou seu objeto de estudo como uma variável dependente formada por cinco variantes. Assim, a indeterminação do sujeito pode ser expressa por uma das seguintes formas:⁹

(i) forma canônica da partícula *se*

(35) *Dá-se* uma banda do dono da terra e uma banda de quem travaia.

(ii) uma das formas do pronome da 1ª pessoa do plural (*nós* ou *a gente*), realizada foneticamente ou não

(36) *É, a gente* procura onde tá a terra mais molhada, *a gente* sabe que dá água mais rápido, né?

(iii) forma do pronome da 2ª pessoa do singular *você*

(37) *Você*, sem sua mãe, ‘*cê* não tem nada na vida.

8. Para uma descrição mais detalhada dessas comunidades, ver Lucchesi *et alia* (2009).

9. A partir de agora os exemplos são retirados do *corpus* analisado, com a transcrição adotada no Projeto *Vertentes* (cf. <http://www.vertentes.ufba.br/projeto/transcricao>).

(iv) forma do pronome da 3ª pessoa do plural, realizada ou não, sem qualquer referência anterior

(38) Chegô lá, *eles* apicaró logo a injeção nela e tal, cortô.

(v) forma verbal não marcada da 3ª pessoa do singular

(39) E, como depois que *prantô* ocalipto que secô bastante água, (...)

Na amostra de fala analisada, foram depreendidas 2.316 formas gramaticais que expressavam a indeterminação do sujeito. O processamento quantitativo dos dados revelou as frequências brutas de cada uma das cinco variantes descritas acima, expressas na tabela abaixo:

Tabela 1: Frequência de uso das formas variantes de indeterminação do sujeito no português afro-brasileiro

| Variante | Nº de ocorrências | Frequência |
|----------------------------------------------|-------------------|-------------|
| Pronome da 1ª pessoa do plural | 1.040 | 45% |
| Forma verbal da 3ª pessoa do singular | 692 | 30% |
| Pronome da 3ª pessoa do plural | 356 | 15% |
| Pronome da 2ª pessoa do singular <i>você</i> | 220 | 9,5% |
| Forma canônica <i>se</i> | 8 | 0,5% |
| TOTAL | 2.316 | 100% |

De imediato, chama atenção a baixíssima frequência da forma canônica *se*: menos de meio por cento do total de ocorrências (mais precisamente, 0,35%). Pode-se dizer que essa forma praticamente não faz parte da gramática das comunidades rurais afro-brasileiras. No outro extremo, observa-se que a forma mais usual de expressar a indeterminação do sujeito é emprego da primeira pessoa do plural, realizada pronominalmente ou não. Essa variante compreende quase a metade das ocorrências do *corpus* analisado: 45% do total. Em seguida vem a forma verbal da 3ª pessoa do singular com sujeito nulo, sem referência anterior, com 30% do total de ocorrências. O pronome da 3ª pessoa do plural, realizado ou não, é a 3ª estratégia mais empregada para expressar a indeterminação do sujeito. Com 15% do total de ocorrências, corresponde à metade da frequência de emprego da forma simplificada da 3ª pessoa do singular e a apenas um terço da frequência do pronome da 1ª pessoa do plural. Entre as formas dos pronomes pessoais, o *você* é a forma menos empregada para a expressão da indeterminação do sujeito, com pouco menos de dez por cento do total de ocorrências (9,5%).

Nas subseções abaixo, buscaremos descrever o encaixamento linguístico e social do fenômeno variável em foco. Considerando que o objetivo central deste artigo é identificar os reflexos do contato entre línguas nessa variedade do português brasileiro, a análise focalizará a variante da 3ª pessoa do singular com sujeito nulo, variante que supostamente decorre de mudança que se originou numa situação de transmissão linguística irregular de tipo leve (cf. seção 1 deste artigo). Dessa forma, na análise dos condicionamentos linguísticos e sociais, as ocorrências das outras quatro variantes serão agrupadas em oposição às ocorrências dessa variante, tornando binária a variável dependente.

5.1. Encaixamento linguístico da simplificação morfológica na forma de expressar a indeterminação do sujeito no português afro-brasileiro

A análise variacionista focalizando a variante da forma verbal da 3ª pessoa do singular com sujeito nulo revelou que o seu emprego é condicionado pelas seguintes variáveis estruturais: (i) o paralelismo discursivo, (ii) a referência do falante a si mesmo e (iii) o tipo de frase. Os resultados de cada uma dessas variáveis serão apresentados a seguir.

5.1.1. A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro segundo a variável *paralelismo discursivo*

O *paralelismo formal* é um princípio que busca dar conta da tendência à repetição de uma opção formal na cadeia da fala (SCHIFFRIN, 1981). Adotamos aqui a formulação de Scherre e Naro (1993), que distingue a repetição da mesma escolha linguística no interior da oração, o chamado *paralelismo formal*, da repetição numa sequência de orações, o chamado *paralelismo discursivo*. Os resultados apresentados na Tabela 2 abaixo se situam no plano do paralelismo discursivo, no qual se considera a forma de indeterminação do sujeito que foi empregada na oração anterior, nos casos de sequências de orações com a mesma referência. Nos casos de orações isoladas, a ocorrência foi classificada como *primeira menção*. Nessa análise que focaliza a variante da simplificação morfológica, espera-se que, nas orações precedidas por outra oração em que a referência foi feita pela forma da 3ª pessoa singular, predomine largamente a escolha por essa mesma variante.

Tabela 2: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função do *paralelismo discursivo* (Nível de Significância: .010)

| Forma anterior | Nº de Oc. / Total | Freq. | P.R. |
|---------------------------------------------|-------------------|-------|------|
| Forma verbal da 3ª pessoa do singular | 447/555 | 80,5% | .933 |
| Primeira menção | 160/527 | 30,5% | .602 |
| Pronome da 3ª pessoa do plural | 25/237 | 10,5% | .285 |
| Pronome da 2ª pessoa do singular <i>ocê</i> | 12/169 | 7% | .229 |
| Pronome da 1ª pessoa do plural | 47/819 | 5,5% | .176 |
| TOTAL | 691/2.307 | 30% | .221 |

Os resultados da tabela acima confirmam claramente o previsto pelo princípio do paralelismo discursivo. Nas orações precedidas por outra oração que contém um sujeito indeterminado expresso pela forma da 3ª pessoa do singular, a frequência de uso dessa variante quase que triplica, passando de 30% de frequência geral para 80,5% nesse contexto. O peso relativo de .933 confirma esse grande favorecimento. Nas orações isoladas, a frequência dessa variante é praticamente idêntica à frequência geral, indicando uma neutralidade desse fator. Entretanto, o peso relativo de .602 indica um pequeno favorecimento dessa variante nesse contexto. Nas orações precedidas por orações com sujeito indeterminado expresso de outra maneira, a frequência dessa variante reduz-se fortemente, o que se confirma com os resultados dos pesos relativos. Nesses casos, a maior frequência da variante morfológicamente simplificada ocorre quando o sujeito indeterminado da oração anterior é expresso pela 3ª pessoa do plural, o que aproxima, em princípio, essas duas variantes.

5.1.2. A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro segundo a variável *referência ao próprio falante*

Estudos anteriores têm revelado que a produção linguística do falante muda nos enunciados em que ele se refere a si mesmo (e.g., LUCCHESI, 2008b). Além disso, o fenômeno em foco está relacionado ao nível de referencialidade do sujeito da oração, o que se relaciona claramente com o fato do próprio falante estar incluído no escopo referencial do sujeito da oração ou não. Nos casos de indeterminação universal, em que o escopo referencial do sujeito abarca potencialmente todos os indivíduos da espécie humana, a referência inclui o falante; nos casos em que a indeterminação do sujeito abarca uma coletividade, o falante pode estar incluído ou não nessa coletividade; e nos casos de referência a um grupo específico, cuja identidade não é revelada no discurso, o falante é necessariamente excluído (cf. seção 2 deste artigo). Não obstante a variável *nível de referencialidade do sujeito indeterminado* não ter sido selecionada como estatisticamente relevante, seus efeitos se manifestam indiretamente na variável aqui em foco, como se pode ver nos resultados da Tabela 3 abaixo:

Tabela 3: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função de referência ao próprio falante (Nível de Significância: .018)

| Referência ao próprio falante | Nº de Oc. / Total | Freq. | P.R. |
|-------------------------------|-------------------|-------|-------------|
| Exclui o falante | 220/576 | 38% | .583 |
| Inclui o falante | 472/1.732 | 27% | .472 |
| TOTAL | 692/2.308 | 30% | .287 |

Os resultados da tabela acima revelam que a frequência da escolha pela forma morfológicamente simplificada de expressar a indeterminação sujeito aumenta quando o escopo referencial desse sujeito indeterminado exclui o falante, com o peso relativo se elevando de um input geral de .287 para .583, nesse contexto. Isso significa também que a forma simplificada é menos usada nos casos de indeterminação universal e de indeterminação parcial, aproximando-se, assim, da variante da 3ª pessoa plural, o que já havia ocorrido na variável anterior do paralelismo discursivo.

5.1.3. A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro segundo a variável *tipo de frase*

A variável *tipo de frase* também se situa no plano discursivo e busca aferir a influência sobre a forma de expressar a indeterminação do sujeito decorrente do fato de o falante estar fazendo uma pergunta, ou uma afirmação contrafactual, em oposição a uma simples declaração afirmativa. Os resultados dessa variável são apresentados na Tabela 4 abaixo:

Tabela 4: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função do tipo de frase (Nível de Significância: .018)

| Tipo de frase | Nº de Oc. / Total | Freq. | P.R. |
|----------------------------|-------------------|-------|-------------|
| Frase Interrogativa | 18/38 | 47,5% | .720 |
| Frase Negativa | 66/212 | 31% | .519 |
| Frase Declarativa | 608/2.058 | 29,5% | .494 |
| TOTAL | 692/2.308 | 30% | .287 |

Os resultados quantitativos, tanto das frequências absolutas, quanto dos pesos relativos, revelam que a variante da 3ª pessoa do singular é favorecida quando o falante está fazendo uma pergunta, ao passo que o efeito dos dois outros contextos é mínimo. As perguntas favorecem a escolha pela forma simplificada de expressar a indeterminação do sujeito, pois a frequência de uso dessa variante sobe, nesse contexto, de 30 para 47,5%, com elevação do peso relativo para .720. Pode-se dizer que a frase negativa favorece muito ligeiramente essa variante, enquanto, na frase declarativa, há uma redução mínima do seu uso. Nas frases negativas, o aumento é de apenas um ponto percentual (peso relativo de .519, bem próximo à neutralidade de .500). Nas frases declarativas, a queda é ainda menor: meio ponto percentual (com peso relativo de .494, igualmente próximo à neutralidade).

5.2. Encaixamento social da simplificação morfológica na forma de expressar a indeterminação do sujeito no português afro-brasileiro

Na chamada abordagem em *tempo aparente*, a configuração encontrada nas variáveis sociais permitiria ao sociolinguista fazer um diagnóstico entre a *variação estável* e a *mudança em progresso* (LABOV, 1972 e 1981). No primeiro caso, o processo de variação estaria estabilizado, ao passo que, no segundo caso, estaria em curso na comunidade de fala uma mudança no sentido da generalização de uma determinada variante, em detrimento de outra(s). E o resultado de tal análise em relação à variação nas formas gramaticais de expressar a indeterminação do sujeito pode fornecer evidências empíricas significativas em favor da hipótese da influência do contato entre língua na formação das comunidades de fala analisadas.

A hipótese a ser testada é a de que a variante da 3ª pessoa do singular com sujeito nulo decorreria da simplificação morfológica prevista nos processos de transmissão linguística irregular de tipo leve que ocorrem em certas situações de contato linguístico massivo. Segundo essa hipótese, essa forma gramatical de indeterminar o sujeito teria predominado na fala das comunidades analisadas no passado. Mas, atualmente, outras formas de indeterminação do sujeito, sobretudo o emprego das formas da 1ª pessoa de plural e da 2ª pessoa do singular, estariam avançando nas comunidades, em função da influência linguística dos grandes centros urbanos sobre todas as regiões do país, em um processo de *nivelamento linguístico*, impulsionado pela ação dos meios de comunicação de massa, pela ampliação do sistema de ensino público e pelo deslocamento populacional, entre outros fatores (LUCCHESI, 2001, 2002 e 2006). Se esse raciocínio estiver correto, deveremos encontrar nas comunidades de fala analisadas um processo de mudança em que o uso da variante simplificada da 3ª pessoa do singular estaria em declínio. Essa mudança em favor de formas mais marcadas morfológicamente seria liderada pelos mais jovens e pelos indivíduos que têm mais contato fora da comunidade e alguma experiência de escolarização.

A variável idade é a base para a análise em tempo aparente (NARO, 2003). Se uma variante é usada com mais frequência na medida em que se passa das faixas etárias mais altas, dos indivíduos mais velhos, para a fala dos mais jovens, tem-se o primeiro indício de uma mudança em progresso em favor dessa variante. O cenário aqui esperado seria o contrário, com o uso declinante da variante morfológicamente simplificada da 3ª pessoa do singular com sujeito nulo, já que ela seria proveniente de mudanças induzidas pelo contato entre línguas que teriam ocorrido na formação das comunidades de fala analisadas. Os resultados dessa variável são apresentados na Tabela 5 abaixo, seguida de sua representação gráfica.

Tabela 5: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função da faixa etária do falante (Nível de Significância: .018)

| Faixa etária | Nº de oc./Total | Frequência | Peso Relativo |
|-----------------|-----------------|------------|---------------|
| 21 a 40 anos | 213/882 | 24% | .439 |
| 41 a 60 anos | 205/744 | 27,5% | .473 |
| mais de 60 anos | 274/682 | 40% | .608 |
| TOTAL | 692/2.308 | 30% | .287 |

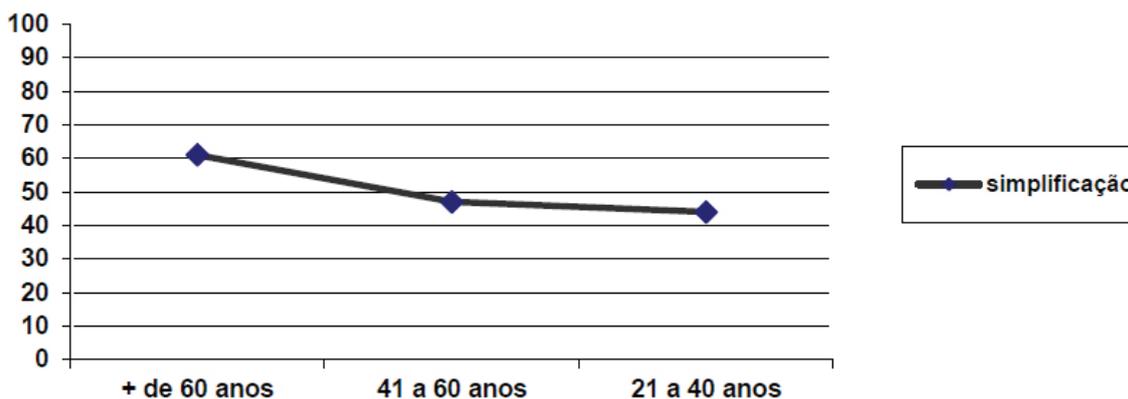


Gráfico 1: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função da faixa etária do falante (em peso relativo)

Os resultados revelam que a variante morfológicamente simplificada é mais usada pelos falantes mais velhos das comunidades (com mais de 60 anos), e a frequência de uso dessa variante cai acentuadamente quando se passa para a faixa etária seguinte (de 41 a 60 anos), passando 40% para 27,5% (P.R. de .608 contra .473). Já a diferença na frequência de uso dessa variante entre essa faixa intermediária e a dos mais jovens (de 21 a 40 anos) é bem menor, passando de 27,5% para 24% (P.R. de .473 contra .439). De qualquer forma, a linha declinante, como representada no Gráfico 1 acima, já é um indício de uma mudança em progresso, com a substituição da forma simplificada por formas morfológicamente mais marcadas de expressão do sujeito indeterminado, o que vai ao encontro da hipótese aqui assumida sobre o processo sócio-histórico de formação das comunidades rurais afro-brasileiras e seus desdobramentos na sincronia atual. Porém, a hipótese de que a variante simplificada teria se originado em processo de transmissão linguística irregular e estaria sendo substituída agora por formas mais marcadas morfológicamente precisa ser confirmada também pelos resultados das demais variáveis sociais.

O perfil sociolinguístico das comunidades rurais afro-brasileiras isoladas tem revelado que os homens tendem a liderar as mudanças em direção às variantes de prestígio (LUCCHESI, 2009b). Isso se explica porque, nessas comunidades, os homens têm mais contato com o mundo exterior, enquanto as mulheres ficam mais confinadas ao universo familiar e da roça. Esse resultado, que encontra paralelo em outros estudos sociolinguísticos do português popular (RODRIGUES, 1992; BORTONI-RICARDO,

2011), contrapõe-se ao resultado geral de análises sociolinguísticas que focalizam variedades urbanas, realizadas em países plenamente industrializados (CHAMBERS, 1995). Os resultados encontrados na presente análise confirmaram o perfil já definido para as comunidades rurais afro-brasileiras, como se pode ver na Tabela 6 abaixo:

Tabela 6: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função do sexo do falante (Nível de Significância: .018)

| Sexo | Nº de oc./Total | Frequência | Peso Relativo |
|------------------|-----------------|------------|---------------|
| Feminino | 424/1.165 | 36,5% | .566 |
| Masculino | 268/1143 | 23,5% | .433 |
| TOTAL | 692/2.308 | 30% | .287 |

Os resultados quantitativos mostram que a variante da simplificação morfológica predomina na fala das mulheres, com frequência de 36,5% (P.R. de .566) contra 23,5% (P.R. de .433) da fala dos homens. Esses últimos, por ter mais contato com o mundo exterior que as mulheres, incorporaram mais as formas mais morfológicamente marcadas de expressão da indeterminação do sujeito, como o uso do *nós/ a gente* e do *você*, devido à influência dos centros urbanos. Assim, vai se constituindo um cenário de mudança linguística de cima para baixo e de fora para dentro das comunidades aqui analisadas. Esse cenário se confirma com os resultados da variável escolaridade, como se pode ver na tabela a seguir:

Tabela 7: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função da escolaridade do falante (Nível de Significância: .017)

| Escolaridade | Nº de oc./Total | Frequência | Peso Relativo |
|-----------------------|-----------------|------------|---------------|
| Analfabeto | 447/1.279 | 35% | .533 |
| Semianalfabeto | 245/1.029 | 24% | .459 |
| TOTAL | 692/2.308 | 30% | .293 |

A variante que supostamente resultaria da influência do contato entre línguas na formação das comunidades é mais frequente na fala dos indivíduos totalmente analfabetos (35%, com P.R. de .533). Já na fala dos indivíduos que têm algum letramento (o que, em alguns casos, se restringe à habilidade de desenhar o nome) é mais frequente o emprego das formas que presumivelmente estariam se fortalecendo nos hábitos linguísticos da comunidade, em função de influências externas.

O diagnóstico de uma mudança de cima para baixo e de fora para dentro tem sua confirmação definitiva com os resultados da variável que distingue os indivíduos que viveram pelo menos seis meses fora da comunidade daqueles que sempre permaneceram no seio de sua comunidade. Esses resultados são apresentados na Tabela 8:

Tabela 8: A simplificação morfológica na expressão da indeterminação do sujeito no português afro-brasileiro, em função da estada fora da comunidade (Nível de Significância: .017)

| Estada fora | Nº de oc./Total | Frequência | Peso Relativo |
|--------------------|------------------------|-------------------|----------------------|
| Não | 478/1464 | 33% | .522 |
| Sim | 214/844 | 25,5% | .461 |
| TOTAL | 692/2.308 | 30% | .293 |

Os falantes que viveram pelo menos seis meses fora da comunidade usam mais as variantes mais marcadas morfológicamente para expressar o sujeito indeterminado do que a variante da 3ª pessoa do singular com sujeito nulo, que é mais frequente na fala dos que nunca viveram fora da comunidade: 33%, com peso relativo de .522 contra 25,5% e .461, dos que já viveram fora da comunidade. Portanto, os resultados de todas as variáveis sociais convergem para tecer um cenário de mudança em progresso, em que a variante simplificada proveniente do contato entre línguas estaria sendo substituída por variantes morfológicamente marcadas, em função de influências externas provenientes dos grandes centros urbanos.

6. CONCLUSÃO

Este artigo apresentou uma análise da simplificação morfológica na expressão gramatical do sujeito indeterminado na fala de comunidades rurais relativamente isoladas e formadas, em sua maioria, por descendentes diretos de africanos escravizados no interior do Brasil. Uma breve descrição dos processos gramaticais usados para expressar a indeterminação do sujeito nas línguas humanas revelou que o emprego do sujeito nulo para expressar o sujeito de referência genérica em línguas crioulas resultaria de um processo de simplificação morfológica, que constitui uma das características mais proeminentes dessas línguas. Com base nessa premissa, o uso da forma verbal 3ª pessoa do singular, com sujeito nulo, fora de uma cadeia de correferência, para expressar o sujeito indeterminado na fala popular brasileira pode ser interpretado como um reflexo de mudanças induzidas pelo contato entre línguas que se inserem na formação histórica dessa variedade linguística, em função do multilinguismo presente nos primeiros séculos da história do Brasil.

A análise em tempo aparente da variação na forma gramatical de expressar a indeterminação do sujeito na fala de quatro comunidades rurais afro-brasileiras do interior do Estado da Bahia revelou que a variante da 3ª pessoa do singular estaria perdendo terreno para outras formas gramaticais de expressar o sujeito de referência genérica, como a 1ª pessoa do plural. Se a variante simplificada teria predominado nas comunidades analisadas em sua formação, devido ao contato entre línguas, estaria ocorrendo atualmente a difusão de formas morfológicamente mais marcadas, em função do processo de *nivelamento linguístico* em curso no país, no qual a norma linguística de prestígio nos grandes centros urbanos, estaria se difundindo para todas as demais regiões, por meio da ação dos meios de comunicação de massa, do deslocamento populacional e da ampliação do sistema de ensino público, entre outros fatores. No plano do encaixamento linguístico, a análise quantitativa revelou que o fenômeno é condicionado precipuamente por fatores discursivos e semânticos.

Se a variante da 3ª pessoa do singular resulta de um processo de simplificação morfológica inerente às situações de contato linguístico massivo, espera-se que ela predomine nas comunidades de fala brasileiras que teriam sido mais diretamente afetadas por esse processo em sua formação, como as comunidades de fala aqui analisadas (algumas delas oriundas de antigos quilombos). Assim, um desdobramento desta pesquisa seria o estudo contrastivo do fenômeno em diferentes variedades do português brasileiro, para testar empiricamente essa hipótese. Outras frentes de investigação que se descortinam a partir de agora são: por um lado, a análise de registros históricos da fala de escravos africanos e crioulos, para esclarecer a gênese dessa construção no português brasileiro, e, por outro lado, a análise da fala popular e rural, em Portugal, para comprovar a ausência dessa construção no chamado português europeu. Os resultados dessas investigações poderão confirmar essa ampliação do efeito do contato entre línguas na configuração atual do português brasileiro.

REFERÊNCIAS

Bickerton, D. (1981). *Roots of language*. Ann Arbor: Karoma.

_____. (1999). How to acquire language without positive evidence: what acquisitionists can learn from Creoles? In: Degraff, M. (ed.). *Language creation and language change: creolization, diachrony, and development*. Cambridge: The MIT Press, 49-74.

Bortoni-Ricardo, S. M. (2011). *Do campo para a cidade: estudo sociolinguístico de migração e redes sociais*. São Paulo: Parábola.

Cavalcante, S. R. (1999). *A indeterminação do sujeito na escrita padrão: a imprensa carioca dos séculos XIX e XX*. Dissertação de Mestrado. Rio de Janeiro: UFRJ.

Chambers, J. (1999). *Sociolinguistic Theory: linguistic variation and its social significance*. Oxford: Blackwell.

Cunha, C. (1976). *Gramática da Língua Portuguesa*. 3ed. Rio de Janeiro: FENAME.

Ferraz, L. I. (1979). *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press.

Field, F. (1997). Second language acquisition in Creole genesis. In G. Escure & A. Schwegler (eds.), *Creoles, contact and language change: linguistic and social implications*. Amsterdam/Philadelphia: John Benjamins, 127-160.

Guy, G. (1981). *Linguistic variation in Brazilian Portuguese: aspects of phonology, syntax and language history*. Tese de Doutorado. Pennsylvania: University of Pennsylvania.

Guy, G. & Zilles, A. (2007). *Sociolinguística quantitativa*. São Paulo: Parábola.

Hagemeyer, T. (2009). As línguas de S. Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, 1:1, 1-27.

Klein, W. & Perdue, C. (1997). The basic variety (or couldn't natural languages be much simpler?). *Second Language Research*, London, 13:4, 301-347.

Labov, W. (1994). *Principles of linguistic change*. Oxford: Cambridge: Blackwell.

_____ (1981). What can be learned about change in progress from synchrony descriptions. In D. Sankoff & Cedergren H. (eds.). *Variation Omnibus*. Carbondale: Linguistic Research, 177-199.

_____ (1972). *Sociolinguistics patterns*. Oxford: Basil Blackwell.

Lefebvre, C. (1998). *Creole genesis and the acquisition of grammar: the case of Haitian Creole*. Cambridge: Cambridge University Press.

Lopes, C. R. (2003). *A inserção de 'a gente' no quadro pronominal do português*. Frankfurt am Main/Madrid: Vervuert/Iberoamericana.

Lucchesi, D. (2013). O contato entre línguas e a origem do português brasileiro. In E. Gugenberger; H. Monteagudo; & G. Rei-Doval (eds.), *Contacto de línguas, hibrididade, cambio: contextos, procesos e consecuencias*. Santiago de Compostela: Consello da Cultura Galega, 191-221.

_____ (2012). *A diferenciação da língua portuguesa no Brasil e o contato entre línguas*. *Estudos de Lingüística Galega*, Santiago de Compostela, n. 4, 45-65.

_____ (2009). História do contato entre línguas no Brasil. In D. Lucchesi; A. Baxter; & I. Ribeiro (eds.), *O português afro-brasileiro*. Salvador: Edufba, 41-73.

_____ (2009). Conclusão. In: Lucchesi; D.; Baxter, A.; & Ribeiro, I. (org.). *O Português Afro-Brasileiro*. Salvador: Edufba, 513-546.

_____ (2008). Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto. In: Roncarati, C.; Abraçado, J. (Org.). *Português Brasileiro II: contato linguístico, heterogeneidade e história*. Niterói: EdUFF, 366-390.

_____ (2008). A concordância nominal em estruturas passivas e de predicativo do sujeito em comunidades rurais afro-brasileiras isoladas no contexto da história sociolinguística do Brasil. In S. Votre & C. Roncarati (eds.), *Anthony Julius Naro e a lingüística no Brasil: uma homenagem acadêmica*. Rio de Janeiro: 7Letras, 148-168.

_____ (2006). Parâmetros sociolinguísticos do português brasileiro. *Revista da ABRALIN* 5:1, 83-112.

_____ (2003). O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In C. Roncarati & J. Abraçado (eds.), *Português brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7Letras, 272-284.

_____ (2002) Grandes territórios desconhecidos. *Linguística (ALFAL)*, São Paulo, 14, 191-222.

_____ (2001). As duas grandes vertentes da história sociolinguística do Brasil. *DELTA*, São Paulo, 17:1, 97-130.

Lucchesi, D. & Baxter, A. (2006). Processos de crioulização na história sociolinguística do Brasil. In S. Cardoso; J. Mota; & R. V. Mattos e Silva (eds.), *Quinhentos anos de história linguística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 163-218.

Lumsden, J. S. (1999). Language acquisition and creolization In M. Degraff (ed.), *Language creation and language change: creolization, diachrony, and development*. Cambridge: The MIT Press, 129-157.

Mattos e Silva, R. V. (2004). *Ensaio para uma sócio-história do português brasileiro*. São Paulo: Parábola.

McWhorter, J. H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology*, 5:2/3, 125-166.

_____ (1998). Identifying the creole prototype: vindicating a typological class. *Language*, 74:4, 788-818.

Naro, A. (2003). O dinamismo das línguas. In M. C. M. Mollica & M. L. Braga (eds.), *Introdução à Sociolinguística*. São Paulo: Contexto, 43-50.

Nunes, J. M. (1990). *O famigerado "se": uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador*. Dissertação de Mestrado. Universidade de Campinas, Campinas.

Rocha-Lima, C. H. (1960). *Gramática Normativa da Língua Portuguesa*. 5ed. Rio de Janeiro: Briguiet & Cia.

Rodrigues, A. (1992). *Língua e contexto sociolinguístico: concordância no português popular de São Paulo*. Araraquara: UNESP.

Rougé, J.-L. (2008). A inexistência de crioulo no Brasil. In J. L. Fiorin & M. Petter (eds.), *África no Brasil: a formação da língua portuguesa*. São Paulo: Contexto, 63-74.

Scherre, M. & Naro, A. (1993). Duas dimensões do paralelismo formal na concordância verbal no português popular do Brasil. *DELTA*, São Paulo, 9:1, 1-14.

Schiffrin, D. (1981). Tense variation in narrative. *Language*, 57:1, 5-62.

Siegel, J. (2008). *The emergence of Pidgin and Creole languages*. Oxford: Oxford University Press.

Wekker, H. (1996). Creolization and the acquisition of English as a second language. In _____ (ed.), *Creole Languages and language acquisition*. Berlin: Mouton de Gruyter, 139-149.