

Análise sociolinguística da variação na concordância nominal de gênero no português indígena sateré-mawé da Amazônia

Sociolinguistic analysis of variation in gender agreement in saterémawé indigenous portuguese from Amazonia

Dante Lucchesi*
Hellen Picanço**

RESUMO

Seguindo os pressupostos da Sociolinguística Variacionista, o artigo apresenta os resultados de uma análise, em tempo aparente, da variação na concordância de gênero no Sintagma Nominal (SN), na variedade de português falado, em situação de bilinguismo, pelo povo indígena sateré-mawé, do Estado do Amazonas. Os dados foram retirados de oito entrevistas informais, com membros da comunidade, e submetidos ao cálculo estatístico, com o programa GoldVarb X. A variação na concordância nominal de gênero, que não ocorre normalmente no português popular do Brasil, resultaria da situação de bilinguismo da comunidade. E os resultados da análise quantitativa revelaram uma situação de mudança em progresso com a implementação do uso da regra de concordância na comunidade. Os fatores estruturais que condicionam o processo de mudança são: a configuração sintagmática do SN, a propriedade de flexão de gênero do nome núcleo do SN e sua vogal temática.

Palavras-chave: concordância de gênero, português indígena, variação linguística, mudança linguística, bilinguismo.

Recebido em 1 de junho de 2020.

Aceito em 5 de novembro de 2020.

DOI: <http://dx.doi.org/10.18364/rc.v1i60.452>

*Universidade Federal Fluminense, dante.lucchesi@gmail.com, orcid.org/0000-0002-8058-2658

**Universidade Federal do Amazonas, india.parintintins@gmail.com, orcid.org/0000-0002-8058-2658

ABSTRACT

Following the assumptions of Variationist Sociolinguistics, the paper presents the results of an analysis, in apparent time, of the variation in gender agreement in the Nominal Phrase (NP), in the variety of Portuguese spoken by the Sateré-Mawé indigenous people, in the State of Amazonas, in a situation of bilingualism. The data were taken from eight informal interviews, with members of the community, and subjected to statistical calculation, with the GoldVarb X program. The variation in nominal gender agreement, which does not normally occur in Brazilian popular Portuguese, would result from the situation of bilingualism in the community. And the results of the quantitative analysis revealed a situation of change in progress with the implementation of the use of the agreement rule in the community. The structural factors conditioning the change process are the syntagmatic configuration of the NP, the gender inflection property of the NP's core noun and its thematic vowel.

Keywords: gender agreement, Indigenous Portuguese, linguistic variation, linguistic change, bilingualism.

Introdução

A variação na concordância de número do Sintagma Nominal (e.g., *as menina bonita* por *as meninas bonitas*) é geral em todas as variedades do português brasileiro (PB), havendo, porém, uma diferença na frequência de uso da regra que separa as variedades populares da elite letrada do país (SCHERRE, 1994). Porém, a variação no que concerne à categoria gramatical de gênero não é normalmente encontrada no português popular brasileiro (PPB), mesmo na zona rural (VEADO, 1982), onde as variedades linguísticas são mais desviantes do padrão normativo. A variação na concordância de gênero no SN (e.g., *um foto velho* por *uma foto velha*) fica, portanto, restrita a variedades do português que passaram por um processo mais intenso de contato entre línguas em sua formação, como o chamado *português afro-brasileiro* (PAB), a linguagem falada em comunidades rurais afro-brasileiras isoladas (LUCCHESI; BAXTER; RIBEIRO, 2009). Essa relação da variação na concordância nominal de gênero com o contato entre línguas massivo e radical é reforçada pelo fato de desse mecanismo ser virtualmente eliminado no processo de criouliização, como o testemunham os crioulos portugueses

da costa ocidental da África (LUCCHESI, 2000; 2009a). A variação na concordância nominal de gênero também se encontra no português vernáculo de Angola, que é uma variedade de português falada como segunda língua (L2) por falantes das diversas línguas banto deste país (INVERNO, 2009), o que reforça ainda mais a relação entre o fenômeno e o contato entre línguas.

Não haveria, no geral, evidências diretas de que a variação na concordância de gênero observadas no PPB seja derivada de mudanças induzidas pelo contato, pois as variedades do PPB, como o PAB, são faladas atualmente por comunidades monolíngues. Naro e Scherre (2007), por exemplo, argumentam que a variação na concordância nominal e verbal no PB seria, essencialmente, o resultado de uma *deriva secular*, que teria suas origens já na passagem do latim ao português e que o contato do português com as línguas indígenas e africanas no Brasil só a teria acelerado. Contudo, uma análise mais atenta da história da morfologia nominal de gênero, na formação da língua portuguesa e em seu desenvolvimento ulterior, revela um processo de fortalecimento e ampliação do rendimento funcional da flexão de gênero, de modo que, ao invés de uma tendência a uma perda gradual das marcas de gênero, o que se vê é uma ampliação da marcação gramatical do gênero no desenvolvimento histórico da língua portuguesa (LUCCHESI, 2000 e 2009a).

Além disso, há uma vasta literatura produzida pela pesquisa linguística sobre situações de contato em geral que comprovam a relação empiricamente motivada entre o contato linguístico massivo e um processo de simplificação morfológica que atinge, sobretudo, as regras de concordância, que não têm valor informacional (WINFORD, 2003; MCWHORTER, 2007; AUTOR; TRUDGILL, 2001, 2009; 2010). Assim, a hipótese de que a atual variação no uso das regras de concordância nominal e verbal nomeadamente nas variedades populares do PB é o resultado de mudanças induzidas pelo contato entre línguas no passado tem-se mostrado muito mais plausível (LUCCHESI, 2008, 2012a, 2013, 2015a), sobretudo diante das fragilidades históricas, teóricas e empíricas da hipótese da deriva, conforme demonstrado por Lucchesi (2012b). De qualquer forma, a observação de como a concordância

de gênero se apresenta em variedades do PB que atualmente se encontram em situação de contato aportaria evidências empíricas significativas em favor da hipótese do contato.

O panorama linguístico do Brasil na atualidade se caracteriza por um monolinguismo amplamente hegemônico, na medida em que mais de 98% dos mais de 210 milhões de brasileiros têm como língua materna apenas o português. Nesse contexto, as situações de contato entre línguas, multilinguismo, bilinguismo e afins, são bem localizadas e marginais, compreendendo: as situações de fronteira, em que o PB convive basicamente com variedades do espanhol sul americano; as comunidades de imigrantes; e os povos indígenas. Entre esses últimos, pode-se visualizar um *continuum* desde povos indígenas que perderam sua língua nativa e são monolíngues em português até povos que preservam seu modo de vida, cultura e língua quase que integralmente, entre os quais a penetração do português é muito restrita. Em um nível intermediário, encontram-se as comunidades indígenas que ainda preservam sua língua nativa, mas experimentam um crescente bilinguismo, em função de sua integração na sociedade branca dominante. Por ser hegemônica e mais viável em termos socioeconômicos, a língua portuguesa tem-se imposto cada vez mais entre os povos indígenas, ameaçando a conservação da maioria das línguas indígenas que são faladas no Brasil (RODRIGUES, 1993).

As variedades de português faladas nas comunidades indígenas bilingues exibem, geralmente, um espectro mais amplo de variação do que aquele observado nas variedades monolíngues do PPB, em função não apenas dos efeitos gerais do contato, como a simplificação morfológica já referida acima, mas também em função do que Weinreich (1953) denominou *interferências* da língua materna sobre a língua a ser adquirida, a língua alvo, ou segunda língua. Essas interferências constituem o que se denomina *influência* ou *transferência do substrato*, quando uma língua dominante se impõem sobre a língua nativa da comunidade, como é o caso das comunidades indígenas, onde o português se impõem sobre a língua indígena, podendo levar

ao um processo de mudança de língua (ing. *language shift*), no qual a língua da comunidade cai em desuso e a língua dominante se impõe plenamente (HICKEY, 2010; ROMAINE, 2010). Diante disso, a descrição e análise das línguas indígenas é da maior relevância, como forma de documentar e conservar a diversidade linguística do país (RODRIGUES, 1993).

Paralelamente, o estudo do chamado *português indígena* é outrossim importante, não apenas pela relevância de se conhecer a diversidade étnica da língua portuguesa no Brasil, mas também pela observação direta dos efeitos do contato entre línguas e do bilinguismo sobre a estrutura linguística do português. Os resultados de tais análises podem lançar luzes, tanto sobre a forma como as situações de contato massivo afetam a estrutura das línguas em geral, quanto sobre a história sociolinguística do Brasil, na qual o português se foi impondo crescentemente sobre comunidades indígenas, levando à substituição e ao desaparecimento da imensa maioria de suas línguas. Com esses propósitos, foi realizada a análise da variação na concordância nominal de gênero no português indígena sateré-mawé, cujos resultados aqui se apresentam.

As comunidades indígenas sateré-mawé, do município de Parintins, no Estado do Amazonas são bilingues, sendo que a maioria dos seus membros, sobretudo nas aldeias, tem a língua sateré-mawé como língua materna e falam o português como segunda língua. Contudo, a penetração do português é cada vez maior, em função das pressões socioeconômicas, sobretudo nas novas gerações e na periferia das cidades, que já abriga parte dessa etnia. Assim, a hipótese que orientou a análise foi a de que a variação na concordância gênero observada nessa variedade linguística resultaria da situação de bilinguismo em que o português é falado como segunda língua. Entretanto, estaria em curso na comunidade um processo de mudança no sentido do incremento do uso da regra de concordância de gênero, em função do domínio crescente da língua portuguesa, sobretudo entre os membros mais jovens da comunidade.

Para testar essa hipótese, foi realizada uma análise em *tempo aparente*, dentro do enquadramento teórico e metodológico da Sociolinguística

Variacionista (WEINREICH; LABOV; HERZOG, 2006[1968]; LABOV, 2008[1972], 1994, 2001a, 2001b, 2010), que se baseou em uma amostra de fala vernácula recolhida em duas aldeias sateré-mawé, por meio de pesquisa de campo realizada no ano de 2018. A análise buscou identificar os condicionamentos linguísticos e extralinguísticos do fenômeno variável, bem como possíveis interferências da língua indígena, com base em uma análise quantitativa dos dados linguísticos extraídos da amostra de fala constituída, com o emprego do programa Goldvarb X de cálculo estatístico (SANKOFF; TAGLIAMONTE; SMITH, 2005). Pretende-se, com essa análise, contribuir, tanto para a compreensão geral da mudança linguística provocada pelo contato entre línguas, quanto para a conhecimento da configuração atual da realidade sociolinguística brasileira e sobretudo de sua formação histórica.

Para cumprir esses objetivos, este artigo contém duas breves seções introdutórias e uma seção central mais extensa, na qual são apresentados os resultados da análise sociolinguística da variação na concordância de gênero no português indígena sateré-mawé. Na primeira seção introdutória, fala-se das contribuições e da relevância do estudo do chamado português indígena. A segunda seção traça um panorama da variação da concordância gênero na língua portuguesa. E na conclusão, é feita uma síntese dos resultados alcançados e apresentadas as implicações e os desdobramentos desta análise.

1. A importância do estudo do português indígena

Quem observa o panorama linguístico atual do Brasil não imagina que sua formação histórica é profundamente marcada pelo contato entre línguas, pois o Brasil é, na atualidade, um dos países de maior homogeneidade linguística do planeta, com mais 98% de sua população sendo monolíngue em português (LUCCHESI, 2015b). Porém, eram faladas no Brasil mais de mil línguas indígenas, quando se iniciou a colonização portuguesa no século XVI (RODRIGUES, 1993), e mais de 200 línguas africanas teriam sido trazidas

para Brasil, em função da escravização de povos africanos, entre 1550 e 1850 (*circa*) (PETTER, 2006). Portanto, a ampla hegemonia da língua portuguesa na atualidade é o resultado de um violento processo de homogeneização linguística que se perpetrou através do extermínio de povos indígenas, da escravidão e de processos degradantes de aculturação forçada (LUCCHESI, 2009b, 2015b, 2019a, 2019b). Como parte desse processo, as línguas indígenas brasileiras foram reduzidas a menos de trezentas na atualidade, e sua sobrevivência deve-se, sobretudo, ao seu isolamento em regiões mais remotas e pouco acessíveis da Amazônia e do Planalto Central. Contudo, a maioria dessas línguas enfrenta, hoje, um grande risco de desaparecer, em função das crescentes pressões da sociedade capitalista brasileira sobre as sociedades indígenas remanescentes. Dessa forma, o violento processo de homogeneização do país, continua em curso na atualidade. Como resultante linguística desse processo, emergem variedades do português que são faldas em muitos casos como segunda língua em comunidades indígenas bilingües, configurando o que se denomina genericamente *português indígena*.

O Censo de 2010 do IBGE computou quase 900 mil indígenas na população brasileira (896,9 mil), que se distribuem por 305 etnias e falam 274 línguas. Contudo, só 37,4% dos indígenas com 5 anos ou mais falam uma língua indígena, e 76,9% falam o português.¹ Por esses números, obtém-se um percentual de 14,3% de bilingüismo indígena, o que corresponde a 128.257 indivíduos bilingües, em língua indígena e língua portuguesa. Porém, o conceito de *português indígena* se aplica tanto aos indígenas bilingües (ou multilingües, no caso dos que falam mais de uma língua indígena), quanto aos que são monolingües em português, independentemente de terem ou não algum conhecimento de sua língua ancestral (MAHER, 1998). Desse modo, o universo do português indígena recobre uma gama muito

1 Fonte: IBGE, Censo 2010. Disponível em: <https://censo2010.ibge.gov.br/noticias-censo?busca=1&id=3&idnoticia=2194&t=censo-2010-poblacao-indigena-896-9-mil-tem-305-etnias-fala-274&view=noticia>. Acesso em 25 de maio de 2020.

diversificada de contextos socio e etnolinguísticos. Apesar disso, o conceito genérico de *português indígena* (PI) se sustenta em função de um conjunto de características socioculturais e linguísticas comuns, que permitem reunir esse espectro de falares em uma variedade mais geral, que é heterogênea, como toda variedade linguística tomada em seu contexto social de uso, como o *português popular brasileiro*, ou mesmo o *português brasileiro*.

Porém, vamos nos concentrar aqui nas variedades de português faladas por grupos indígenas bilingues ou multilíngues. Com esse escopo, o PI se define, no plano sociocultural, por ser falado por grupos geralmente pouco numerosos, que falam uma língua ou mais línguas indígenas, que é ou são minoritárias, em face do português, língua dominante em termos socioeconômicos. Esses grupos vivem em graus variados de isolamento, sendo geralmente o extrativismo e/ou a agricultura de subsistência sua principal atividade econômica. Tal definição engendra uma série de parâmetros socioeconômicos variáveis, que devem ser considerados na análise sociolinguística de cada variedade de português indígena: (i) sua representatividade numérica; (ii) seu grau de isolamento/integração em relação à sociedade dominante; (iii) suas atividades econômicas. Esses parâmetros socioeconômicos, por sua vez, determinam parâmetros culturais e linguísticos que devem ser analisados e mensurados: (i) vitalidade da identidade étnica; (ii) conservação da cultura material e imaterial do povo indígena; (iii) conservação de sua língua; (iv) penetração do português na comunidade.

Todos esses parâmetros estão interligados. Quanto maior o isolamento, maior a conservação do *modus vivendi* e da cultura ancestral do povo, maior a vitalidade de seu sentimento de identidade e de sua língua, e menor a penetração do português. No outro extremo, temos os povos muito integrados, aculturados, com baixo sentimento de pertencimento étnico, com a língua desaparecida ou em vias de desaparecimento, tendo como língua hegemônica o português. Entre os dois extremos há todo um gradiente complexo e diversificado, pois não se pode pensar em termos de uma relação

mecanicista entre os fatores objetivos e subjetivos, uma vez que um povo indígena não precisa estar isolado para conservar sua cultura e valorizar sua identidade étnica e *a fortiori* sua língua nativa (FOUGHT, 2010).

Sobretudo nas últimas décadas, tem havido uma reação, com movimentos de conservação e valorização da cultura e da identidade indígena, até com ações de restauração linguística (MAIA, 2006; ALMEIDA; RUBIM, 2012; D'ANGELIS, 2014). Porém, trata-se de uma reação proporcionalmente diminuta, diante de séculos de violência e dominação econômica, simbólica e cultural, que levou ao extermínio de povos inteiros, ou de processos violentos e devastadores de aculturação forçada, de modo que integração à sociedade dominante significa para os povos indígenas, no mais das vezes, exploração, espoliação, violência física e/ou simbólica, destruição do seu habitat, aculturação forçada, perda da identidade e da língua. A redução de mais de mil línguas indígenas faladas no território brasileiro, quando se iniciou a colonização portuguesa, no século XVI, para as menos de trezentas línguas atuais nada mais é do que reflexo dramático de um genocídio secular e de um processo forçado e devastador de aculturação, que, aviltando os padrões de comportamento do indígena, relativamente à sexualidade, ao trabalho, à indumentária, à religião e à língua, conduziam-no igualmente à sua desagregação identitária, à depressão e à morte (FREYRE, 2002).

E o fato de a maioria das línguas indígenas hoje correr sério risco de extinção significa que todo esse processo ainda está em curso, embora tenha sido um pouco mitigado e processos e ações de resistência devam ser registrados pelo seu grande valor e significado, embora ainda muito reduzidas e limitadas. Assim, sendo, o quadro geral do contato do português com as línguas indígenas deve ser definido como uma situação em que uma língua dominante em termos socioeconômico e demográfico se impõe sobre uma língua minoritária, de modo que o português vai substituindo a língua indígena, progressivamente, nas diversas esferas de uso da língua, até o desaparecimento desta (HICKEY, 2010; ROMAINE, 2010). A base socioeconômica desse processo é a desintegração do *modus vivendi* e da

cultura ancestral do povo indígena e a integração dos seus membros no mercado de trabalho e de consumo da sociedade dominante.

Do ponto de vista das funções de uso da língua, o português penetra como língua de comunicação com mundo exterior. Em seguida, passa a ser a língua de uso dos indivíduos que se integram no comércio ou nas atividades produtivas com a sociedade branca, bem como das crianças que se alfabetizam em português. Muitas vezes, os adultos passam a falar em português com os filhos ou incentivar neles o uso dessa língua, como meio para que eles possam se inserir melhor na realidade socioeconômica que se impõe. Assim, na situação mais comum, a língua indígena vai se restringindo em seu espectro funcional e a proficiência nela vai decaindo, à medida que se passa para as gerações mais jovens, de modo que, no mais das vezes, os mais jovens só tem um conhecimento passivo ou dominam um reduzido vocabulário da língua dos seus ancestrais.²

Portanto, a situação de cada variedade do português indígena vai variar bastante, consoante o grau de integração do povo que a fala, por um lado, e grau de conservação cultural e linguística desse povo, por outro. O português pode ser, desde a língua já hegemônica, ou mesmo única do povo indígena, até ser uma segunda língua adquirida e usada em graus variados de proficiência, em um espectro funcional igualmente variado; ou seja, podendo ser essa variedade bem restrita, em seu grau de proficiência e espectro funcional. Quanto mais situado nesse último polo, maiores serão os efeitos do contato linguístico sobre essa variedade de português, sendo mais notáveis mesmo os processos de interferência ou de transferência da língua indígena para a língua portuguesa.³

2 Esboça-se aqui apenas um quadro muito geral e simplificado, que não considera, por exemplo, as situações de multilinguismo indígena, com casamentos interétnicos e/ou adoção de uma língua de outro povo indígena no qual um povo minoritário se integra, bem como os efeitos da miscigenação com os brancos.

3 Para uma discussão sobre a diferença entre interferência e transferência, veja-se Winford (2003).

Conquanto recubra um grande espectro de variedades, o PI deve ser caracterizado como uma variedade do português popular brasileiro, já que as comunidades indígenas geralmente se situam na base da pirâmide social do país. Desse modo, as características gerais do PPB, como, por exemplo, a ampla variação na concordância nominal e verbal, devem-se reproduzir no PI. Além disso, as variedades do PI devem exibir características linguísticas específicas decorrentes da situação de bilinguismo ou multilinguismo em que são faladas, que podem estar inserido em um processo de mudança de língua (ing. *language shift*) para o português. Nas situações de bilinguismo ou multilinguismo e de mudança de língua, em que os falantes de uma língua minoritária adquirem uma língua dominante como segunda língua (L2), em situações naturais de aquisição não tutorada, há sempre simplificação da L2, sobretudo no que diz respeito ao uso de afixos flexionais e regras de concordância nominal e verbal. E esses processos de simplificação são determinados pela falta de convergência entre a L1 e a L2, podendo ocorrer também interferências ou transferências da L1 sobre a L2 (WINFORD, 2003; TRUDGILL, 2009, 2010; MATRAS, 2010). Portanto, a variação na concordância nominal de gênero que se observa no português indígena saterémawé deve ser vista como o reflexo da situação de bilinguismo em que essa variedade do PB é falada, o que se coaduna com a fato de que, ao contrário do que se observa com a concordância nominal de número, a concordância de gênero é restrita a algumas variedades mais periféricas do PB.

2. A variação na concordância de gênero em português

Quem analisa o panorama sociolinguístico do Brasil, identificando os fenômenos em variação no nível da morfossintaxe, depara-se com uma questão instigante: por que a variação na concordância nominal de número é tão geral, enquanto a variação na concordância nominal de gênero é um fenômeno bem restrito?

No plano formal da gramática, a expressão da categoria gramatical de número, na língua portuguesa, é regular e tem um significado uniforme, atingindo quase a totalidade dos nomes e dos determinantes e modificadores nominais. A flexão de número é feita com o acréscimo do morfema de plural *-s* (com o alomorfe *-es*), sendo o singular não marcado: *casa* : *casas*; *flor* : *flores* etc. São poucos os nomes que não aceitam a flexão de número, mantendo a mesma forma no singular e no plural: *o lápis* : *os lápis*; *o vírus* : *os vírus* etc. Também entre os adjetivos são poucos os que não se flexionam em número: *proposta simples* : *propostas simples* etc. A maioria se flexiona em número, tanto os adjetivos de tema em *-o* quanto os de tema em *-e*: *feio* : *feios*; *inteligente* : *inteligentes* etc; havendo também a alomorfia em *-es*: *montês* : *monteses* etc. Já os determinantes do nome se flexionam sem exceção: *o* : *os*; *esse* : *esses*; *meu* : *meus*; *algum* : *alguns* etc.⁴ No plano semântico, o significado dessa flexão é bem geral e preciso, opondo ‘um’ a ‘mais de um’. Ainda no plano do significado, a oposição singular e plural diz respeito apenas aos nomes, e não aos seus determinantes e modificadores, que assumem a forma de plural juntamente com nome núcleo do SN, em função apenas do mecanismo sintático da concordância.

Toda a simplicidade e coerência da marcação do plural nos nomes poderia ser vista, em princípio, como uma barreira à erosão do mecanismo da concordância de número. Porém, não é isso que se observa, e a variação na concordância nominal de número é geral em todas as variedades do português brasileiro (PB), havendo apenas uma diferença na frequência em que ela ocorre que separa os falares populares da fala da elite letrada (SCHERRE, 1994). Assim, a variação na concordância nominal de número no PB estaria historicamente correlacionada ao maciço contato entre línguas que marca a história sociolinguística do Brasil (cf. Introdução).

4 Os quantificadores pronominais vêm quase sempre no plural: *muitos problemas*, *poucas soluções*, *vários palpites* etc.

A categoria gramatical do gênero, por sua vez, é, na maioria dos casos, um mero classificador gramatical, sem qualquer implicação semântica. Não há qualquer razão semântica para *cama* ser feminino, e *sofá*, masculino. Nem para *planeta* ser masculino e *linguagem* ser feminino – inclusive, no português arcaico era contrário, neste caso. Essa flutuação na fixação do gênero de algumas palavras, que se observa na história do português, também se observa no desenvolvimento diferenciado das línguas românicas. Assim, *o mar* é masculino em português, mas feminino em francês (*la mer*), *a flor* é feminino em português, mas é masculino em italiano (*il fiore*), e *o leite* é masculino em português, mas feminino em espanhol (*la leche*), podendo-se enumerar muitos outros exemplos. Há, porém, um reduzido conjunto de nomes em que o gênero não é arbitrário, os nomes que se referem a seres animados. Uma parte desses se flexiona em gênero, por meio do morfema de feminino *-a*: *o gato* : *a gata*; *o professor* : *a professora*; *o mestre* : *a mestra*; *peru* : *perua* etc.⁵ Em outra parte deles, o gênero é marcado pela derivação (e.g., *imperador* : *imperatriz*; *galo* : *galinha*), ou lexicamente: *homem* : *mulher*; *boi* : *vaca* etc. Mas há nomes, denominados *comuns de dois gêneros*, em que a oposição masculino/feminino é marcada apenas pelos determinantes e/ou modificadores: *o artista* : *a artista*; *o estudante* : *a estudante* etc. Em todos esses casos, a diferença no gênero gramatical têm valor referencial, refletindo a oposição no chamado gênero natural.⁶ Há ainda, os *sobrecomuns*, que só têm um gênero, podendo se referir tanto a seres do sexo masculino, quanto do feminino: *o cônjuge*; *a criança* etc. Por fim, há os chamados *epicenos*, que só têm um gênero, mas aos quais se agrega os termos *macho* e *fêmea* para indicar o gênero natural: *um jacaré fêmea*; *uma cobra macho* etc.

5 Há uns poucos alomorfes, como *avô* : *avó*, e marca subsidiária na alternância vocálica, como em: *porco* : *porca*.

6 Há, porém, nomes que não se referem a seres animados, em que uma pelo menos aparente flexão de gênero assume significados diversos do de gênero natural, como em: *barco* : *barca*; *porto* : *porta*; *ovo* : *ova*.

Como a maioria dos nomes não se flexiona em gênero, a indicação do gênero se faz pela concordância: *uma foto antiga; um dilema muito profundo* etc. E muitos falantes de português L2 têm dificuldade em marcar corretamente o gênero, sobretudo aqueles cujas línguas maternas não têm marcação gramatical de gênero. Porém, a variação na concordância de gênero, longe de ser geral, mesmo no português popular brasileiro, está circunscrita a algumas poucas variedades. Essa vitalidade da marcação gramatical do gênero pode ser explicada pelo desenvolvimento desse mecanismo morfossintático na história da língua portuguesa desde suas origens latinas.

Embora a língua latina exibisse uma divisão tripartida do gênero gramatical, com um gênero neutro recobrindo o nome dos seres inanimados, muitos nomes que deveriam ser do gênero neutro se distribuíam entre o masculino (e.g., *pees* ‘pé’, *rivus* ‘regato’ e *mensis* ‘mês’) e o feminino (e.g., *mensa* ‘mesa’, *manus* ‘mão’ e *memoria* ‘memória’). Havia uma concentração de palavras masculinas na segunda declinação, enquanto as palavras femininas predominavam na primeira declinação, o que permitia formar *lupa* ‘loba’ em oposição a *lupus* ‘lobo’. Porém, essa oposição entre masculino e feminino ficava diluída na declinação nominal latina de seis casos, com formas distintas para o singular e o plural. Por outro lado, havia uma numerosa classe de adjetivos que, a exemplo de *justus*, *justa*, *justum*, se flexionavam regularmente quanto aos três gêneros gramaticais. E já no período latino, ocorre um aumento da produtividade funcional desse paradigma, que indicava morficamente a oposição masculino/feminino, com a incorporação de adjetivos da chamada 2ª classe, que originalmente não faziam a distinção mórfica entre o masculino e o feminino (SILVA NETO, 1988, p. 231). Esse processo já prenunciava um fortalecimento da oposição morfológica entre o masculino e o feminino. O desaparecimento formal do gênero neutro (incorporado maciçamente ao gênero masculino) e a perda da flexão casual dos nomes fortaleceram sobremaneira a oposição mórfica entre masculino e o feminino, na passagem do latim às línguas românicas.

Assim, o morfema de feminino *-a* se fixa na gramática do português e das demais línguas românicas. Contudo, seu rendimento funcional era muito limitado no português arcaico, com palavras como *senhor*, *pecador*, *pastor* e *burguês* sendo usadas como feminino, sem se flexionar. No século XIV, registram-se na mesma situação: *enganador*, *merecedor*, *vencedor* etc. (MALER, 1964, p. 27). Com a evolução da língua, todas essas formas passaram a admitir regularmente a flexão com o morfema de feminino *-a*: *senhora*, *pecadora*, *pastora*, *burguesa*, *enganadora*, *merecedora*, *vencedora*. E o emprego do morfema de feminino *-a* generaliza-se a partir do século XVI, atingindo os nomes e adjetivos terminados em *-or*, *-ol*, *-ês* e *-nte*, como em: *doutora*, *espanhola*, *portuguesa* e *infanta*. Portanto, o rendimento da flexão de gênero só se incrementa, ao longo da história do português, o que pode explicar o fato de não haver uma ampla variação na concordância de gênero no PPB, como se observa na concordância nominal de número. Diante desse cenário, não cabe, no caso da concordância nominal de gênero, a hipótese de uma deriva secular no sentido da perda gradual de marcas morfológicas. E a relação histórica entre a variação na concordância nominal de gênero e o contato entre línguas se fortalece ainda mais, diante do fato de esse fenômeno só ser observado, ou em variedades do português que passaram por um intenso contato em sua formação, ou em variedades do português que ainda são usadas em situação de bilinguismo ou multilinguismo.

Tal é o caso da variação na concordância de gênero que ocorre no português vernáculo de Angola, a variedade de português falada como L2, por falantes das diversas línguas banto deste país. Inverno (2009, p. 97) informa que, não obstante seja geral entre os falantes jovens e instruídos, “no discurso dos falantes mais velhos e menos instruídos, raramente ocorre concordância de gênero entre o núcleo do SN e os seus determinantes”, como em: *os palavra*, *esses visita* e *o mamã*. Dessa observação inicial, pode-se concluir que a concordância de gênero se vai fixando, à medida que se aprofunda o domínio da língua portuguesa, falada como L2, entre os mais jovens.

No Brasil, Lucchesi (2000; 2009a) fez uma análise variacionista do fenômeno na fala de uma comunidade rural afro-brasileira, denominada Helvécia, que se situa no extremo sul do Estado da Bahia e pode ter passado por um processo de crioulização em sua formação (FERREIRA, 1984). Embora o nível de variação fosse baixo (cerca de 5%), foi suficiente para configurar um caso de variação estruturada. Os achados de Lucchesi são importantes para estudo que se apresenta aqui, pois ele adotou o mesmo enquadramento variacionista. No plano dos condicionamentos estruturais, Lucchesi constatou que:

- (i) a regra de concordância de gênero (CG) é mais aplicada nos SNs com estrutura mais simples, formados apenas por um nome regido por um determinante;
- (ii) adjetivos e modificadores pós nominais são bastante refratários à CG;
- (iii) o quantificador *tudo*, empregado como pluralizador nominal invariável, bloqueia a CG no SN;
- (iv) os núcleos nominais com propriedade de flexão são os mais favoráveis ao emprego da CG no SN;
- (v) entre os nomes femininos sem flexão de gênero, os de tema em *-a* (e.g. *casa, perna*) são os mais favoráveis à CG, e os de tema em *-o* e *-ão* (e.g. *foto, paixão*), os seus maiores obstáculos.

No plano do encaixamento social, Lucchesi observou que a regra de CG é aplicada quase que categoricamente na fala dos mais jovens, de modo que o fenômeno, que tem suas raízes no processo de transmissão linguística irregular produzido pelo contato entre línguas no passado, estaria desaparecendo na atualidade.

A variação na CG foi analisada também no chamado *português de contato* do Parque Nacional do Xingu, uma realidade etnolinguística que se aproxima mais da realidade observada aqui. Na região do Alto Xingu, são faladas pelo menos nove línguas indígenas: kamayurá e aweti, da família

tupi-guarani; wará, mehinaku e yawalapiti, da família aruak; kalapalo, kuikuro e matipu, da família karib; e o trumai, uma língua isolada. Na parte setentrional do parque, ainda se encontram os seguintes povos: suyá, juruna, kayabí, txukahamãe, krenakarore e txicão. Ao todo vivem no Parque quinze grupos indígenas, que formam uma sociedade mais ampla, congregada por objetivos comuns e “um complexo sistema de relações cerimoniais, sociais e econômicas”. Porém, todos esses povos mantiveram suas línguas nativas e têm um grau de contato com o português variável. Assim, a variedade de português que surge a partir do contato Expedição Roncador-Xingu com os indígenas tornou-se a língua franca do Parque, servindo, tanto para a comunicação entre brancos e indígenas, quanto entre os diversos povos indígenas que falam línguas mutuamente ininteligíveis entre si (EMMERICH, 1991).

No início da década de 1980, Charlotte Emmerich constituiu uma amostra linguística do português de contato do parque do Xingu, que ela definiu como uma variedade pidginizada do português, compreendendo um *continuum* que se estende desde variedades bem restritas de português faladas como L2 pelos índios monolíngues até a “competência praticamente bilingue dos jovens do Posto Leonardo Villas Boas, passando por uma gama de situações linguísticas estruturalmente diferenciadas”, segundo o grau dos processos de variação que tradicionalmente são associados à pidginização, como a redução nas flexões nominais de gênero e número e verbais de número e pessoa.

Lucchesi e Macedo (1997) analisaram a variação na CG na amostra constituída por Emmerich, com uma abordagem semelhante à empregada aqui. No plano dos condicionamentos estruturais, a análise variacionista com suporte estatístico constatou que os SNs mais simples, formados apenas pelo determinante e o nome núcleo eram a estrutura que mais favorecia a aplicação da regra de CG. Aplicação da regra também foi favorecida por SNs cujo núcleo era um nome com flexão de gênero (e.g., *professora*, *menina*). Esses resultados estão compatíveis com os encontrados na análise de Lucchesi em Helvécia (cf. supra). Já em relação aos SNs cujos núcleos eram nomes sem flexão de gênero, os resultados foram divergentes. A concordância de gênero

não foi favorecida pelos nomes de tema em *-a*, mas pelos nomes de tema em *-e* e *-*e*, como em *gente* e *vez*, respectivamente. Os autores explicaram isso pelo fato de muitas línguas do Parque possuírem uma grande quantidade de nome de masculinos terminados em *-a*, como *caraíba* ‘homem branco’, que se mantêm inclusive no português de contato.

Portanto, essa análise da variação na CG em uma variedade importante do português indígena revelou interferências das línguas indígenas do substrato no quadro de variação encontrado. Diante disso, a análise aqui desenvolvida não deixará de considerar as características da língua indígena de substrato na análise variacionista do fenômeno.

3. Análise sociolinguística da variação na concordância de gênero do português indígena sateré-mawé

Os Sateré-Mawé habitam a Terra Indígena (TI) Andirá-Marau, localizada na região do Médio rio Amazonas, entre os Estados do Amazonas e do Pará, cujo espaço territorial compreende 788.528 hectares, com perímetro de 477,7 km. O território indígena Sateré-Mawé divide-se em três grandes regiões: rio Andirá, em Barreirinha; rio Waikurapá, em Parintins; e rio Marau-Urupadi, em Maués. A região do Andirá pertence ao município de Barreirinha, onde ficam mais de 50 aldeias, com uma população de aproximadamente 3,8 mil pessoas. Na região do Marau-Urupadi, pertencente ao município de Maués, ficam aproximadamente 37 aldeias, com cerca de 3,3 mil sateré-mawé (TEIXEIRA, 2005), e na região do Uaikurapá, pertencente ao município de Parintins, são apenas sete comunidades: São Francisco, com 172 habitantes; Nova Alegria, com 173 habitantes; Vila da Paz, com 140 habitantes; Vila Batista, com 331 habitantes; Nova Galileia, com 76 habitantes; Monte Carmelo, com 75 habitantes; e Ipiranga, com 110 habitantes (FUNAI/PARINTINS, 2018). A FUNAI (2018) estima que cerca de 519 Sateré-Mawé residem na área urbana de Parintins, sendo que aproximadamente 100 desses indígenas residem fixamente na casa de trânsito da cidade, situada na Rua

Silva Campos, centro, Parintins, AM. O restante dos sateré-mawé vivem em casas localizadas nos bairros periféricos da cidade, sobrevivendo como mão de obra barata no mercado de trabalho informal.

No que concerne às origens da língua sateré-mawé, há duas hipóteses concorrentes. Para Rodrigues e Dietrich (1997), existiria um ramo composto por mawé, aweti e o proto tupi-Guarani, dentro do tronco tupi. Porém, ocorreu uma primeira mudança: a família Mawé se separou do ramo, depois houve a separação entre aweti e o proto tupi-Guarani, formando a configuração atual de famílias linguísticas do tronco tupi, a saber: arikén, aweti, juruna, munduruku, mawé, mondé, puroborá, ramarána, tupari e tupi-guarani. A segunda hipótese é sustentada por Drude (2006), para quem não houve mudanças linguísticas significativas entre mawé, aweti e o proto tupi-Guarani, de modo que as separações entre elas foram simultâneas ou muito próximas.

Na morfologia nominal da língua sateré-mawé, não há flexão de gênero. A distinção de gênero entre pessoas e nomes de parentesco é feita lexicalmente, como em: *ihainia* ‘homem’ e *hariporia* ‘mulher’; *iwot* ‘pai’ *iti* ‘mãe’. Com os nomes que se referem a seres animados, o gênero pode ser indicado pela composição com as palavras *wari’i* (‘mulher/fêmea’) e *pa’iat* (‘homem/macho’), para indicar o gênero feminino e masculino, respectivamente, como em: *moi wari’i* ‘cobra fêmea’ e *moi pa’iat* ‘cobra’; *waipaka pa’iat* ‘galo’ e *waipaka* ‘galinha’; *aware* ‘cachorro’ *aware wari’i* ‘cadela’. Mas já se registram também influências do português nessa área da gramática, como em *hamiariru* ‘neto’ e *hamiarira* ‘neta’, em que se toma de empréstimo a flexão de gênero do português (SILVA, 2010, p. 154-156).

Dentro do enquadramento teórico-metodológico da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006[1968]; LABOV, 2008[1972], 1994, 2001a, 2001b, 2010), o universo da pesquisa é a comunidade de fala sateré-mawé, constituída por 13.310 falantes da língua sateré-mawé (IBGE, 2010). Grande parte desses falantes é bilíngue em sua língua e português. A maior proporção de falantes da língua indígena encontra-se nas aldeias: 95,9% falam sateré-mawé em contexto de terra indígena, enquanto

apenas 54,9% falam sateré-mawé no contexto urbano. Não há diferença linguísticas entre as mulheres e os homens; entre aquelas, 95,9% falam a língua indígena, enquanto o percentual de falantes da língua indígena entre os homens é de 96,1%. No entanto, quando se considera o uso da língua portuguesa pelos sexos, a situação linguística muda bastante: 72,2% dos homens falam português, enquanto apenas 54,4% falam essa língua (TEIXEIRA, 2005).

A análise sociolinguística da variação na concordância de gênero no Sintagma Nominal, no português indígena sateré-mawé, foi feita em amostra linguística constituída em 2018, nas comunidades sateré-mawé de Nova Alegria e Vila Batista, localizadas no rio Waicurapá, município de Parintins, Estado do Amazonas. Foram analisadas oito entrevistas de tipo sociolinguístico, em que o entrevistador busca criar um clima de informalidade, para superar o que Labov (2008[1972]) denominou *paradoxo do observador*. Cada entrevista teve a duração aproximada de 45 minutos, e os entrevistados se distribuíram pelos dois sexos e por três faixas etárias.

Foi feita uma recolha exaustiva de todas as ocorrências de SNs femininos nas entrevistas transcritas. Essa base de dados foi codificada de acordo com as variáveis explanatórias definidas na análise, e essa codificação foi submetida ao programa de cálculo multivariado GoldVarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005), que seleciona as variáveis independentes estatisticamente relevantes e mensura a interferência de cada fator sobre o fenômeno variável em foco, ponderando a ação simultânea de todos os demais fatores relevantes.

Há duas formas de tratar da variação da concordância no SN (SCHERRE, 1988; LUCCHESI, 2000). Na *abordagem sintagmática*, observa-se a aplicação da regra do SN como um todo, enquanto na *abordagem mórfica*, se observa a marcação em cada constituinte do SN.⁷ O mecanismo morfossintático da concordância se aplica no SN, quando todos os elementos flexionáveis do sintagma portam as mesmas marcas referentes às categorias gramaticais de número e/ou gênero. No caso do gênero, todos os elementos

7 Scherre (1988) denomina esta *atomística* e aquela *não-atomística*.

flexionáveis em um SN feminino devem receber o morfema de feminino *-a*, como em: *aquela gata preta e uma foto velha*. Se apenas um elemento do SN deixar de portar a marca de feminino, a concordância não ocorre, como em: *a gata preto e um foto velha*. Quando o núcleo do SN é um nome flexionável em gênero, a marca pode vir somente nesse nome núcleo, como em: *o gata preto*. Mas, no caso dos nomes não flexionáveis em gênero, o SN pode figurar sem qualquer marca de feminino, como em: *um foto velho*.

Assim, a variável dependente desta análise foi definida em termos binários, com o valor positivo de aplicação da regra de concordância, quando todos os constituintes flexionáveis do SN portam a marca de feminino, e com o valor negativo de não aplicação da regra, se ao menos um constituinte flexionável deixou de receber o morfema de feminino. Na amostra de fala analisada, foram depreendidas 1.069 ocorrências de SNs femininos, sendo que a regra de concordância de gênero foi aplicada totalmente em 951 ocorrências, o que representa uma frequência geral de aplicação da regra de 89% do total, como se pode ver na Tabela 1:

Tabela 1. Frequência da concordância de gênero no SN, no português indígena sateré-mawé.

Aplicação da regra de concordância	Nº de ocorrências/Total	Frequência
sim	951/1.069	89%
não	118/1.069	11%

Labov (2003) fixa em cinco por cento a frequência mínima para que ocorra a variação na estrutura linguística. O percentual de onze por cento encontrado permitiu assim observar um processo de variação encaixado na estrutura linguística e social da comunidade de fala analisada. Na primeira subseção a seguir, serão apresentados os resultados da análise quantitativa das variáveis estruturais que afetaram o emprego da regra de concordância gênero no SN como um todo. Na segunda subseção, serão apresentados os resultados quantitativos dos condicionamentos sociais.

3.1. A variação na concordância de gênero no Sintagma Nominal no português indígena sateré-mawé: o encaixamento estrutural

Para aprender os condicionamentos estruturais do fenômeno variável em tela, foram formalizadas sete variáveis independentes ou explanatórias: (i) a distinção entre SNs preposicionados e não preposicionados; (ii) a configuração sintagmática do SN; (iii) a caracterização morfológica do nome núcleo do SN; (iv) a vogal temática do nome núcleo do SN; (v) a concordância de número no SN; (vi) o número absoluto de constituintes do SN; e (vii) o número de constituintes flexionáveis no SN. Desses grupos de fatores, o GoldVarb X selecionou como tendo valor estatístico os grupos: (i) a configuração sintagmática do SN; (ii) a caracterização morfológica do nome núcleo do SN e (iii) a vogal temática do nome núcleo do SN. Além de identificar as variáveis com valor estático (i.é., aquelas que, de acordo com a distribuição quantitativa dos dados, efetivamente influenciaram o fenômeno variável em tela), o programa de regras variáveis fornece as frequências brutas de ocorrência do fenômeno em cada fator considerado em termos percentuais, bem como o *peso relativo* (P.R.), índice resultante do cálculo multivariado que computa a influência simultânea de todos os fatores conjuntamente.⁸ Os valores obtidos para cada variável explanatória selecionada pelo programa serão apresentados e analisados a seguir.

8 Esse índice resulta, portanto, mais confiável que a frequência bruta, pois, em cada ocorrência, todos os fatores atuam simultaneamente, de modo que, por exemplo, se a configuração sintagmática favorece a aplicação da regra de concordância de gênero em uma ocorrência, a natureza de seu nome núcleo pode desfavorecê-la, e assim por diante. No cálculo do peso relativo o Programa pondera a ação simultânea de todas as variáveis independentes.

3.2. A aplicação da regra de concordância de gênero no Sintagma Nominal no português indígena sateré-mawé, segundo a configuração sintagmática do SN

Para aferir como a configuração sintagmática do SN afeta a aplicação da regra de concordância de gênero (CG), essa variável explanatória foi estruturada da seguinte maneira:⁹

- 1) SN formado apenas pelo determinante e um nome núcleo [Det + N]
Ex.: *uma comunidade, a história, esses tecnologia* etc.¹⁰
- 2) SN formado pelo determinante, um nome núcleo e um constituinte que não participa da concordância de gênero (um adjetivo que não se flexiona em gênero, um sintagma preposicionado ou uma oração relativa) [Det+N+SX]
Ex.: *as crianças doentes, a casa dele, essa entrevista que nós tamo tendo, essa união de falá um com o outro* etc.
- 3) SN com um possessivo antes do nome núcleo [...Poss+N...]
Ex.: *o meu carteirinha, nosso língua, minha casa* etc.
- 4) SN com o quantificador *tudo/todo* antes do nome núcleo [...Q...N...]
Ex.: *tudo tapioca, tudo as coisa, todas as filhas que eu tive* etc
- 5) SN com outro modificador antes do nome núcleo [...Mod...N...]
Ex.: *algumas bananas, a primeira vista, uma boa pergunta* etc
- 6) SN com um adjetivo depois do nome núcleo [...N...Adj...]
Ex.: *a irmã mais nova, as coisa errado* etc
- 7) SN com qualquer modificador depois do nome núcleo [...N...Mod...]
Ex.: *essa banda todo, a cultura nossa, coisa nenhuma* etc

9 Alguns fatores inicialmente especificados foram reunidos depois, em função do baixo número de ocorrências.

10 A partir daqui todos os exemplos serão retirados da base de dados da análise.

A hipótese primeira seria a de que os SNs com estrutura mais simples e menos constituintes participando do processo de CG (ou seja, a primeira e segunda configuração acima) exibiriam uma frequência maior de aplicação da regra, como foi observado nas análises de Lucchesi (2000 e 2009a) e Lucchesi e Macedo (1997), para o PAB de Helvécia e o português indígena do Xingu, respectivamente (cf. seção anterior). A presença de um outro elemento participando do processo de CG faria cair a frequência de aplicação da regra, sobretudo se esse outro elemento vier depois do nome núcleo, porque a relação que une o nome núcleo aos constituintes à sua esquerda é mais estreita do que a que o une aos elementos da margem direita do SN. A especificação da presença do quantificador *tudo/todo* à esquerda do núcleo do SN se deve ao fato desse quantificador muitas vezes funcionar como um pluralizador invariável, nessa e em outras variedades do português que se formam em situação de contato massivo,¹¹ o que bloqueia a aplicação da regra de CG, como se pode ver nos exemplos *tudo tapioca* e *tudo as coisa*.¹² Assim configurada, esta variável apresentou os seguintes resultados quantitativos:

Tabela 2. Aplicação da regra de concordância de gênero no SN no português indígena sateré-mawé, em função da configuração sintagmática do SN.

Configuração do SN	Nº de ocorrências/Total	Frequência	Peso relativo
Det + N	558/620	90%	.529
Det+N+SX	154/169	91%	.547
[...Poss+N...]	95/103	92%	.554
[...Mod...N...]	93/110	84,5%	.381
[...Q...N...]	14/20	70%	.160

continua

11 Esse uso também foi observado por Lucchesi (2000) no português afro-brasileiro da comunidade de Helvécia, no extremo sul do Estado da Bahia.

12 Não foi possível isolar o efeito desse constituinte à direita do nome núcleo, em função do baixo número de ocorrências desse tipo na amostra analisada.

Configuração do SN	Nº de ocorrências/Total	Frequência	Peso relativo
[...N...Adj...]	32/39	82%	.308
[...N...Mod...]	04/07	57%	.219
TOTAL	950/1068	89%	

Nível de significância: .043.

Os pesos relativos devem, em princípio, refletir os valores percentuais, sendo que qualquer oscilação nessa relação deve ser considerada em função da interferência de fatores de outras variáveis, o que não aconteceu neste caso, em um nível que seja digno de nota. A leitura dos resultados dos pesos relativos deve ser feita na comparação entre eles, porém os valores próximos a .500 devem ser lidos como de fatores cuja influência sobre o fenômeno analisado deve ser neutra. Valores acima desse patamar indicam um fator que favorece a aplicação da regra, enquanto valores abaixo desse valor desfavorecem-na. Assim, os SNs de configuração mais simples, formados por um determinante e o nome núcleo, com ou sem a presença de um terceiro constituinte que não participa do processo de CG, são os contextos que favorecem ligeiramente a aplicação da regra, com pesos relativos um pouco acima de .500, o que corresponde aos resultados percentuais, já que a frequência de aplicação da regra nesses casos eleva-se ligeiramente, passando de uma frequência geral de 89% para uma frequência de 90%, nos casos em que o SN é formado apenas pelo nome e o determinante (peso relativo de .529), e para 91% (peso relativo de .547), nos casos em que o SN é formado pelo nome, o determinante e um constituinte que não participa do processo de concordância. Nesse contexto, o possessivo se comporta como o artigo e o demonstrativo, ou seja, como um determinante *stricto sensu*, não como um modificador, pois, quando o núcleo é precedido por um possessivo, a frequência de aplicação da regra é de 92%, com peso relativo de .554.

Porém, se o nome núcleo é precedido por um pronome indefinido, um numeral ou por um adjetivo, a frequência de aplicação da regra cai para

84,5%, o que se reflete no peso relativo de .381. Se esse constituinte antes do nome for o quantificador *tudo/todo*, a aplicação da regra cai para 70%, o que corresponde a um peso relativo de .160; ou seja, é o fator que mais inibe a aplicação da regra de CG no SN. A presença desse quantificador, ou de qualquer outro modificador, à direita do nome núcleo, é o segundo fator que mais desfavoreceu a aplicação da regra, com uma frequência de 57% e peso relativo de .219.

Portanto, os resultados quantitativos da variável configuração sintagmática do SN possibilitam duas generalizações. A primeira é a de que a aplicação da regra de CG é favorecida nas estruturas mais simples, o que pode ser definido como *princípio da simplicidade*. A segunda é a de que a presença de um constituinte que se flexiona em gênero na margem direita do SN desfavorece fortemente a CG, já que a coesão entre esses elementos e o núcleo do SN é menor do que a que existe entre o núcleo e os elementos que o precedem, no que pode ser definido como *princípio da coesão*.

3.3. A aplicação da regra de concordância de gênero no Sintagma Nominal no português indígena sateré-mawé, segundo a caracterização morfológica do nome núcleo do SN

Além da constituição sintagmática, a caracterização morfológica do nome núcleo do SN relativamente à categoria gramatical do gênero mostrou-se igualmente um fator relevante no desencadeamento do mecanismo sintático da CG. A fundamentação teórica para a formulação deste grupo de fatores pressupõe que o mecanismo da CG se desencadeia a partir do nome núcleo, já que a categoria gramatical do gênero seria inerente à categoria lexical dos nomes. Desse modo, assumiu-se a existência de uma correlação entre a aplicação da regra de CG e o modo como o gênero é morficamente marcado nos nomes, distinguindo-se, basicamente, os nomes com flexão de gênero (e.g., *sogro/sogra*; *gato/gata*; *mestre/mestra*) *vis-à-vis* os nomes sem flexão de gênero (e.g., *planta*, *mente*, *foto*).

A hipótese gerada por essa formulação teórica é a de que os nomes com flexão de gênero favorecem a CG, enquanto os nomes que não se flexionam desfavorecem-na, como já observado em estudos anteriores sobre o tema (cf. seção anterior). Os resultados quantitativos confirmaram essa hipótese, como se pode ver na Tabela 3:

Tabela 3. Aplicação da regra de concordância de gênero, no português indígena saterémawé, em função da caracterização morfológica do nome núcleo do SN.

Natureza do núcleo do SN	Nº de ocorrências/Total	Frequência	Peso relativo
Nome com flexão de gênero	68/72	94,4%	.739
Nome sem flexão de gênero	883/997	88,6%	.481
TOTAL	951/1069	89%	

Nível de significância: .043.

Quando o nome núcleo do SN se flexiona quanto ao gênero, a frequência de aplicação da regra de concordância sobe de 89% para 94,4%, ao passo que essa frequência cai ligeiramente quando o nome núcleo não se flexiona em gênero. Os pesos relativos confirmaram essas oscilações percentuais. O peso relativo de .739 confirma que o nome com flexão de gênero é fator que favorece bastante a aplicação da regra de CG. E o peso relativo de .481 indica que o nome sem flexão de gênero desfavorece ligeiramente o emprego dessa regra. Retornando ao plano geral da teorização sobre o SN, pode-se pensar que a marcação do gênero parte do nome núcleo para os determinantes da margem esquerda e para os modificadores da margem direita, com proeminência dos primeiros.

Portanto, a conclusão que se pode extrair da análise deste grupo de fatores é: quanto mais explícita for a marcação do gênero no núcleo do SN, maior será a frequência de aplicação da regra de concordância, no que pode ser definido como *princípio da saliência*.

3.4. A aplicação da regra de concordância de gênero no Sintagma Nominal no português indígena sateré-mawé, segundo a vogal temática do nome núcleo do SN

Seguindo o princípio teórico da proeminência do nome núcleo no desencadeamento do mecanismo da CG no interior do SN, foi verificado se a vogal temática dos nomes sem flexão de gênero teria algum efeito sobre o emprego da regra de concordância. A hipótese adotada foi a de que os nomes de tema em *-a* (e.g., *floresta, sola, regra*) favoreceriam a aplicação da regra de CG, em função da homonímia entre essa vogal temática e o morfema de feminino *-a*. Ao contrário os nomes de tema *-o* ou terminados em *-ão*, como *foto* e *comissão*, desfavoreceriam a aplicação da regra, já que a maioria dos nomes masculinos terminam em *-o* ou em *-ão*. Já as palavras de tema em *-e*, como *lente* ou *liberdade*, ou de tema em um **e* teórico, como *mulher* ou *voz*, deveriam se situar em um nível de neutralidade.¹³ Os resultados quantitativos obtidos nesta variável foram os seguintes:¹⁴

Tabela 4. Aplicação da regra de concordância de gênero no SN no português indígena sateré-mawé, em função da vogal temática do nome núcleo do SN.

Vogal temática	Nº de ocorrências/Total	Frequência	Peso relativo
-a	533/617	86,4%	.529
-e	171/177	96,6%	.547
*e	138/152	90,8%	.554
-o ou -ão	57/67	85,1%	.381
TOTAL	899/1013	88,7%	

Nível de significância: .043.

13 Adotamos aqui a descrição proposta por Mattoso Câmara Jr. (1987, p. 87-96).

14 Para aferir essa variável, foram retirados os SNs cujos núcleo fossem nomes com flexão de gênero (e.g., *menina, cachorra, gata* etc). Isso explica o menor número de ocorrências no total desse grupo de fatores.

Os resultados quantitativos confirmaram apenas em parte a hipótese inicial. Com efeito, os nomes terminados em *-o* e *-ão* se mostraram um fator inibidor da aplicação da regra de concordância de gênero, já que a frequência de aplicação da regra caiu de 88,7% para 85,1% nesse contexto. E o peso relativo de .381 indica claramente esse fator como desfavorecedor da aplicação da regra. Porém, os nomes de vogal temática em *-a* não foram o principal fator que favorecesse a concordância de gênero no SN. O peso relativo desse fator (.529) ficou um pouco acima da neutralidade, embora a frequência de aplicação da regra nesse contexto tenha caído ligeiramente, em relação à frequência geral. Os fatores que mais favoreceram a aplicação da regra de concordância de gênero foram os nomes de tema em *-e* e *-*e* teórico, com pesos relativos de .547 e .554, respectivamente.

De qualquer forma, deve-se salientar que o número de ocorrências de cada fator deste grupo reforça o raciocínio que fundamentou a hipótese inicial. Os nomes de tema em *-a* correspondem a 60,9% do total de nomes sem flexão de gênero. No polo oposto, os nomes terminados em *-o* ou em *-ão* perfazem apenas 6,6% do total. Já os nomes de tema *-e* e **e* teórico se situam em um nível intermediário, com 17,5% e 15%, respectivamente. Isso reforça a relação entre a terminação *-a* e o feminino, podendo-se pressupor uma pressão paradigmática em favor dos nomes de tema em *-a*.

Os efeitos dessa variável foram observados na comunidade afro-brasileira de Helvécia, que é monolíngue em português, porém não se confirmaram no português indígena do Parque do Xingu, em função de interferências das línguas do substrato indígena (cf. seção anterior). A mesma interferência do substrato indígena pode ser invocada para explicar o que também se observa aqui. Como já foi dito no início desta seção, não há flexão de gênero na língua sateré-mawé, porém o gênero pode ser indicado pela composição com as palavras *wari'i* (mulher/fêmea) para indicar o gênero feminino, e *pa'iat* (homem/macho) para indicar o masculino. Como se pode ver, a palavra que indica o masculino em sateré-mawé tem, em sua sílaba final, a vogal *a*, o que pode atenuar a pressão paradigmática que existe em

português no sentido de associar essa vogal ao feminino. Por outro lado, a palavra que marca feminino termina com uma vogal alta central *ɨ* e uma vogal alta anterior *i*, sons que se assemelham ao *i* átono, que é a forma como se pronuncia em português a vogal final das palavras terminadas em *-e*, como *rede* (esse *i* átono também aparece na sílaba final do plural das palavras que têm como vogal temática um *-*e* teórico, como em *mulheres* e *vozes*). Essa semelhança pode explicar por que os falantes do português sateré-mawé têm mais facilidade para marcar o gênero das palavras femininas terminadas em *-e* e *-*e* do que as palavras terminadas em *-a* (e.g., *mata*, *roda*, *comida* etc.). Portanto, na frequência como os falantes do português sateré-mawé fazem a concordância gênero em SNs cujos núcleos nominais não são palavras que se flexionam em gênero, manifesta-se uma influência do substrato indígena, como ocorreu no português de contato do Xingu. Isso reforça a ideia de que o português indígena constitui uma variedade diferenciada do português popular do Brasil, devido às interferências provenientes do substrato indígena.

4. A variação na concordância de gênero no Sintagma Nominal no português indígena sateré-mawé: os condicionamentos sociais

Um dos principais objetivos da análise sociolinguística em *tempo aparente* é determinar se um processo de variação observado sincronicamente reflete um processo de *variação estável* ou um processo de *mudança em progresso* (LABOV, 2008[1972], 1981, 1994, 2001a). No primeiro caso, as formas variantes tenderiam a permanecer em variação na comunidade indefinidamente. No segundo caso, observa-se a tendência de uma variante ser substituída pela outra. Para fundamentar esse diagnóstico, o sociolinguista baseia-se nos resultados quantitativos de variáveis sociais, como: idade, sexo, classe social ou nível de escolaridade. A idade é a principal variável na análise em tempo aparente, sendo o incremento do emprego da variante inovadora,

na medida em que se passa dos membros mais velhos para os membros mais novos da comunidade, a base para a identificação de um processo de mudança em progresso. O diagnóstico relativamente às demais variáveis sociais, vai variar em função da composição e estruturação de cada comunidade de fala analisada (LUCCHESI, 2004 e 2015a).

No que concerne ao fenômeno aqui analisado, a variação na concordância de gênero nas comunidades do povo indígena sateré-mawé decorre da situação de bilinguismo em que o português é adquirido como segunda língua ou simultaneamente à língua sateré-mawé. Em situações de contato entre línguas como essa, há a tendência à ocorrência de simplificação morfológica da língua que penetra na comunidade, no caso o português (WINFORD, 2003; MCWHORTER, 2007; AUTOR; TRUDGILL, 2001, 2009; 2010). Esse processo costuma afetar mais os mecanismos gramaticais sem valor informacional, como é caso dos mecanismos de concordância nominal e verbal. Por outro lado, observa-se a tendência nas comunidades indígenas brasileiras da penetração cada vez mais forte do português, já que as condições socioeconômicas e culturais são cada vez mais adversas à conservação das línguas indígenas. A assimilação socioeconômica e cultural dos povos indígenas pela sociedade capitalista brasileira faz com que as crianças e os jovens tenham mais motivação em adquirir português do que a língua ancestral dos seus antepassados. Assim, a tendência é a de que a proficiência em português seja maior entre os mais jovens do que entre os mais velhos.

Portanto, em relação ao fenômeno variável observado, o que se espera é que a variação no emprego da regra de concordância gênero, entre os sateré-mawé, seja menor entre os membros mais jovens da comunidade do que entre os mais velhos, configurando um cenário de mudança em progresso, no qual o emprego da regra de concordância tende a se fixar na gramática da comunidade. Para além da variável faixa etária, também foram controladas nesta análise as variáveis sexo e nível de escolaridade do informante.

4.1. A aplicação da regra de concordância de gênero no Sintagma Nominal no português indígena sateré-mawé, segundo a idade do falante

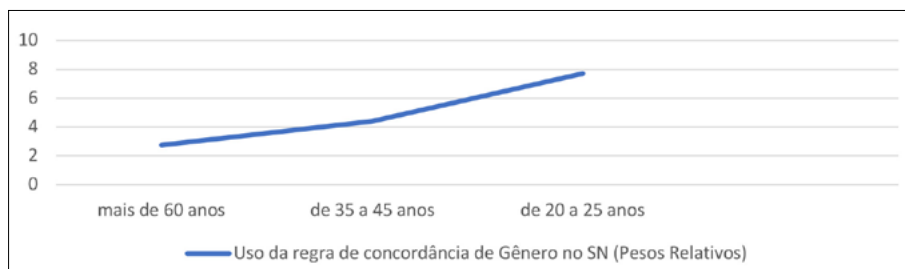
De acordo com a hipótese de que estaria ocorrendo nas comunidades indígenas sateré-mawé uma mudança no sentido da implementação da regra de concordância gênero no SN, o resultado da variável idade do informante deveria apresentar uma curva ascendente em que a frequência de aplicação da regra aumentaria na medida em que se passasse das faixas dos falantes mais velhos para os mais novos. Os resultados dessa variável são apresentados na tabela 5 e podem ser visualizados no gráfico 1:

Tabela 5. Aplicação da regra de concordância de gênero no SN no português indígena sateré-mawé, em função da idade do falante

Idade	Nº de ocorrências/Total	Frequência	Peso relativo
de 20 a 25 anos	289/298	97%	.771
de 35 a 45 anos	463/527	87,9%	.441
mais de 60 anos	1968/243	81,5%	.275
TOTAL	950/1068	89%	

Nível de significância: .043.

Gráfico 1. Efeito da variável faixa etária sobre o uso da regra de concordância de gênero no português sateré-mawé



Os resultados da variável faixa etária confirmam a hipótese de mudança em progresso no sentido da implementação da regra de concordância gênero na comunidade, devendo-se destacar que essa variável foi a primeira a ser selecionada com valor estatístico pelo GoldVarb X. Os falantes com mais de 60 são os que menos empregam a regra, com uma frequência de 81,5%, enquanto os mais jovens exibem a maior frequência, 97%, ficando os falantes de 35 a 45 anos em um nível intermediário, mas um pouco mais próximo dos mais velhos. Isso fica mais nítido nos pesos relativos, que destacam os jovens como o fator que mais favorece a aplicação da regra, com peso relativo de .775, enquanto os pesos relativos das duas faixas dos falantes mais velhos ficam a abaixo da neutralidade, sendo que os falantes com mais de 65 são o fator que mais inibe a aplicação da regra, com peso relativo de .275, mais do que os falantes de 35 a 45 anos, que tem o peso relativo de .441. No gráfico 1, fica clara a curva ascendente que se acentua na passagem da faixa entre 35 e 45 anos para a faixa entre 20 e 25 anos, revelando que o ritmo da mudança tenha aumentado nas duas últimas décadas.

Portanto, pode-se concluir, em primeiro lugar, que a mudança está se acelerando no passado recente da comunidade, o que se explica pela maior integração dos membros mais jovens da comunidade na sociedade branca dominante, com maior escolarização e mais influência dos meios de comunicação de massa e da Internet. Em segundo lugar, deve-se salientar que a frequência de aplicação da regra de 97%, entre os falantes de 20 a 25 anos, situa-se no plano que Labov (2003) definiu como de regra semicategórica, ou seja, o uso da regra de concordância de gênero no SN está se generalizando no português indígena sateré-mawé, como ocorre nas demais variedades do português popular no Brasil. Assim, esse aspecto diferenciador do português indígena, decorrente da situação de bilinguismo, estaria desaparecendo no caso dos sateré-mawé. Os resultados das outras duas variáveis sociais também apontaram para uma situação de mudança em progresso.

4.2. A aplicação da regra de concordância de gênero no Sintagma Nominal no português indígena sateré-mawé, segundo o sexo do falante

Tradicionalmente, as pesquisas sociolinguísticas indicavam que as mulheres tendiam a liderar as mudanças em direção as variantes do padrão de maior prestígio dentro da comunidade (CHAMBERS, 1995, p. 102-103), porém essa generalização tem sido questionada nos últimos tempos (LUCCHESI, 2015a; FREITAG, 2015). Como enfatiza Lucchesi (2015a), o efeito de cada variável social deve ser definido no contexto sócio-histórico específico de cada comunidade de fala analisada, não se podendo isolar um efeito universal de uma variável como sexo/gênero. Assim, o papel da mulher no processo de mudança linguística deverá ser definido pela situação social da mulher em cada comunidade social específica.

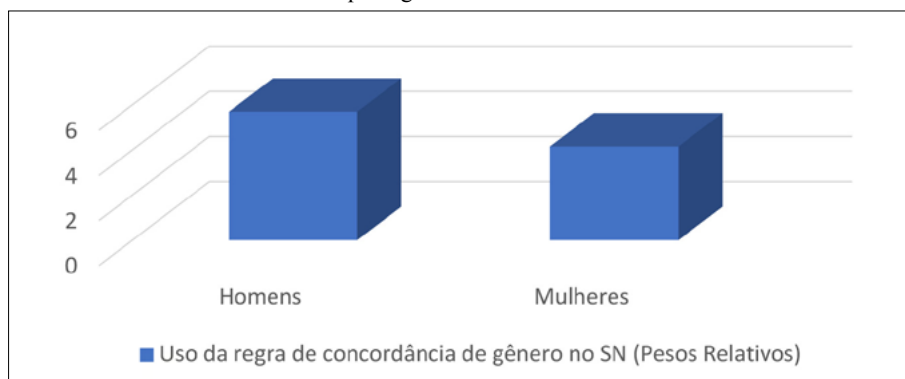
No caso das classes populares do Brasil, são os homens que tendem a liderar as mudanças de cima para baixo, nas quais as variantes do padrão de maior prestígio nos grandes centros urbanos estão sendo assimiladas pelos segmentos populares, tanto na periferia das grandes cidades, quanto nas comunidades rurais (LUCCHESI, 2009c, 2015a). Isso se deve ao fato de os homens, em média, terem mais contato com o mundo exterior e estarem mais bem inseridos no mercado de trabalho do que as mulheres nessas comunidades populares. Dessa forma, no caso aqui estudado, a expectativa seria a de que os homens liderassem o processo de mudança no sentido da implementação da regra de concordância gênero nas comunidades sateré-mawé. Os resultados do processamento quantitativo dos dados referentes à variável sexo são apresentados na Tabela 6 e podem ser visualizados no Gráfico 2:

Tabela 6. Aplicação da regra de concordância de gênero no SN no português indígena sateré-mawé, em função do sexo do falante.

Sexo	Nº de ocorrências/Total	Frequência	Peso relativo
Homens	531/591	89,8%	.569
Mulheres	419/477	87,8%	.415
TOTAL	950/1068	89%	

Nível de significância: .043.

Gráfico 2. Efeito da variável sexo sobre o uso da regra de concordância de gênero no português sateré-mawé



Os resultados quantitativos confirmaram a hipótese inicial. Os homens aplicam a regra de concordância de gênero em uma frequência um pouco maior do que as mulheres (89,8% contra 87,8%), mas os pesos relativos, que resultam de um cálculo multivariado, no qual o efeito de cada fator é mensurado considerando o efeito simultâneo de todos os demais fatores, são claros, indicando a diferença em favor da liderança dos homens no processo de mudança, com .569 para os homens contra .415 para as mulheres. Portanto, assim como acontece em outras comunidades populares do Brasil, os homens tendem a liderar, também nas comunidades indígenas, as mudanças em que as variantes linguísticas hegemônicas são assimiladas pela comunidade. Isso

ocorre porque, em média, os homens têm mais contato com o mundo exterior e têm uma melhor inserção no mercado de trabalho.

4.3. A aplicação da regra de concordância de gênero no Sintagma Nominal no português indígena sateré-mawé, segundo o nível de escolaridade do falante

Um processo de mudança de fora para dentro da comunidade em que uma variante de prestígio se impõe, corresponde ao que Labov (2008 [1972], 2001) denomina *mudança de cima para baixo*. Essa expressão de cima para baixo se refere, tanto à estrutura social, quanto ao nível de consciência, pois os falantes, ao tomarem consciência do prestígio social da variante, fazem um esforço consciente para empregá-la. Nesse contexto os falantes com um nível maior escolaridade tendem a liderar o processo de mudança, já que uma das funções desempenhadas pela escola é difundir a norma padrão, enfatizando as formas que são valorizadas socialmente e condenando o uso de formas sofrem uma avaliação negativa explícita (VOTRE, 2007). Assim sendo, a expectativa seria a de que os falantes indígenas com maior nível de escolaridade liderassem a mudança no sentido de implementação da regra de concordância gênero no SN. Os resultados dessa variável social são apresentados na Tabela 7¹⁵ e podem ser visualizados no Gráfico 3:

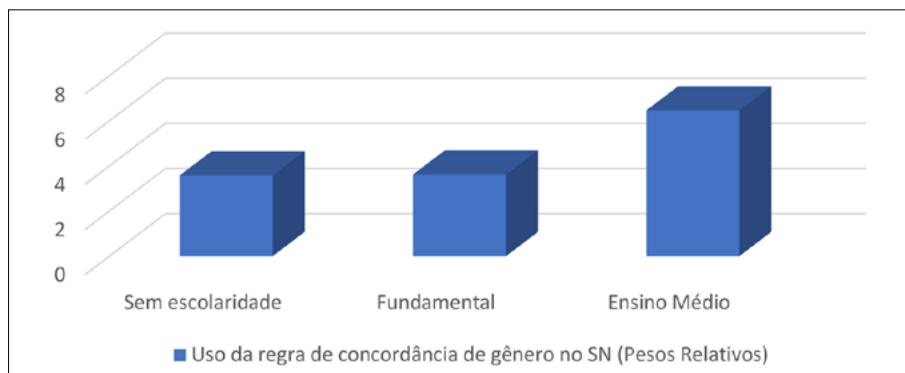
15 Na primeira rodada do GoldVarb, essa variável não foi selecionada como estatisticamente significativa. Isso se deveu a um problema na distribuição da amostra de fala analisada, pois os falantes mais jovens eram igualmente os que tinham um maior nível de escolaridade. Essa superposição fez com que, ao cruzar a variável idade, que foi a primeira a ser selecionada, com o nível de escolaridade, o Programa descartasse esta última, já que o Programa “interpreta”, nesses casos, que o fator determinante é o primeiro a ser selecionado. Portanto, o procedimento adotado foi retirar a variável idade e fazer uma nova rodada. Nessa segunda rodada, o nível de escolaridade foi a primeira variável a ser selecionada, mantendo-se a seleção das outras mesmas variáveis linguísticas e sociais. Deve-se notar também que nível de significância nessa segunda rodada foi ainda melhor do que o da primeira.

Tabela 7. Aplicação da regra de concordância de gênero no SN no português indígena sateré-mawé, em função do sexo do falante.

Escolaridade	Nº de ocorrências/Total	Frequência	Peso relativo
Sem escolaridade	164/199	82,4%	.357
Fundamental	292/344	84,9%	.361
Médio	494/525	94,1%	.645
TOTAL	950/1068	89%	

Nível de significância: .031.

Gráfico 3. Efeito da variável sexo sobre o uso da regra de concordância de gênero no português sateré-mawé



Os resultados do processamento quantitativo confirmaram o efeito do fator escolaridade no processo de implementação da regra de CG na comunidade, pois quanto maior é o nível de escolaridade do indivíduo, maior é a frequência de aplicação da regra. Os indivíduos sem escolaridade empregam a regra de CG em uma frequência de 82,4%, os indivíduos do nível fundamental empregam a regra com uma frequência de 84,9%, e os de nível médio de escolaridade, com uma frequência de 94,1%. Observe-se que a diferença é maior em favor dos indivíduos do Ensino Médio. Os pesos relativos

confirmam isso, o fator Ensino Médio tem o peso relativo de .645, indicando que só esse nível de escolaridade favorece efetivamente a aplicação da regra; já os pesos relativos dos indivíduos do nível fundamental e dos indivíduos sem escolaridade são, respectivamente, .361 e .357, indicando, por um lado, que não há muita diferença entre eles e que esses fatores desfavorecem a aplicação da regra de concordância de gênero. Portanto, os resultados quantitativos apontam no sentido de que a escolaridade favorece a aquisição da regra de concordância de gênero na comunidade indígena em estudo, mas esse fator só atua efetivamente quando o indivíduo chega ao Ensino Médio.

Conclusão

A análise sociolinguística de base quantitativa da variação na concordância de gênero no interior do SN na variedade de português das comunidades indígenas sateré-mawé do município de Parintins, no Estado do Amazonas, revelaram um processo de implementação da regra de nessas comunidades. Esse processo de mudança em progresso é liderado pelos indivíduos mais jovens (na faixa de 20 a 25 anos), do sexo masculino e que cursaram o Ensino Médio. A mudança se encontra em um estágio avançado de implementação, já que o nível de frequência de emprego da regra entre os indivíduos mais jovens é da ordem de 97%, o que já configura uma situação de regra semicategórica (LABOV, 2003).

O encaixamento estrutural do processo de mudança revelou que a aplicação da regra de CG é mais favorecida quando o SN é constituído apenas por um determinante e o núcleo nominal, o qual pode vir ou não seguido por um constituinte que não participa da concordância de gênero (um sintagma preposicionado, um adjetivo que não se flexiona em gênero ou uma oração relativa). A frequência de aplicação da regra de CG aumenta ligeiramente se o determinante for um possessivo. Por outro lado, outros modificadores, como os pronomes indefinidos, e sobretudo o quantificador *todo/tudo/toda*,

que podem figurar antes do núcleo nominal, desfavorecem a CG; da mesma forma, os modificadores à direita do núcleo também se mostraram um fator que desfavorece a CG no interior do SN. Além da configuração sintagmática, a natureza do nome núcleo se revelou, outrossim, um fator que interfere na aplicação do mecanismo da CG no SN, o que reforça a visão de que a marcação do gênero no SN se desencadeia a partir do nome núcleo. Dessa forma, os nomes que se flexionam em gênero favorecem a CG, enquanto os nomes que não se flexionam em gênero desfavorecem-na. A análise aqui desenvolvida definiu, então, que a aplicação da regra de CG é regida pelos princípios da simplicidade, coesão e saliência.

Por outro lado, entre os nomes que não se flexionam em gênero, os nomes femininos terminados em *-o* e *-ão* desfavorecem a concordância de gênero, em decorrência de sua similitude como os nomes masculinos, enquanto os nomes de tema em *-e* e **e* teórico, como *liberdade* e *vez*, são os que mais favorecerem a concordância de gênero, seguidos dos nomes de tema em *-a*, como *corda*. Esse último resultado, que foge ao padrão esperado, em função da homonímia entre a vogal temática *-a* e o morfema de feminino em português, se explica pela influência do substrato indígena, em função da semelhança das vogais finais da palavra *wari'i* ‘mulher/fêmea’, que é usada para indicar o gênero feminino na língua sateré-mawé, e a vogal temática *-e* ou **e*, que se realiza como um *i* átono final em português. Esse padrão divergente reflete, portanto, uma interferência do substrato indígena que individualiza essa variedade do português brasileiro.

A análise de outros processos de variação que também individualizam essa variedade do português brasileiro poderão ampliar ainda mais o conhecimento da diversidade etnolinguística da língua portuguesa no Brasil, bem como a compreensão de como o contato entre línguas afeta a gramática das línguas em geral e da língua portuguesa em particular.

Referências

ALMEIDA, Alfredo Wagner Berno de; RUBIM, Altaci Corrêa. Kokama: a reconquista da língua e as novas fronteiras políticas. **Revista Brasileira de Linguística Antropológica**, v. 4, n. 1, p. 67-80, 2012

CARDOSO, Denise Porto. **Atitudes linguísticas e avaliações subjetivas de alguns dialetos brasileiros**. São Paulo: Blucher, 2015.

CHAMBERS, Jack. **Sociolinguistic theory: linguistic variation and its social significance**. Oxford: Blackwell, 1995.

D'ANGELIS, Wilmar. Línguas Indígenas no Brasil: urgência de ações para que sobrevivam. In: BOMFIM, A. B.; COSTA, F. V. F. da (orgs). **Revitalização de língua indígena e educação escolar indígena inclusiva**. Salvador: Egba, 2014. p. 93-117.

DRUDE, S. On the position of the Awetí language in the Tupi family. In: DIETRICH, W; SYMEONIDIS, H. (Eds). **Guaraní y “Mawetí-Tupí-Guaraní**. Berlin: Lit Verlag, 2006.

EMMERICH, Charlotte. O português de contato no parque indígena do Xingu. **Revista Internacional de Língua Portuguesa**. Lisboa: v. 5/6, p. 37-49, dez. 1991.

FERREIRA, Carlota. Remanescentes de um falar crioulo brasileiro. In: FERREIRA, Carlota et al. **Diversidade do português do Brasil**. Salvador: EDUFBA, p.21-32, 1984.

FOUGHT, Carmen. Ethnic Identity and Linguistic Contact. In: HICKEY, R. (ed.). **The Handbook of Language Contact**. Chichester: Wiley-Blackwell, p. 282-298, 2010.

FREITAG, Raquel. (Re)discutindo sexo/gênero na sociolinguística. In: FREITAG, R. M. K., SEVERO; GORSKI, C. (ed.). **Mulheres, Linguagem e Poder - Estudos de Gênero na Sociolinguística Brasileira**. São Paulo: Editora Edgard Blücher, p. 17-74, 2015.

FREYRE, Gilberto. Casa Grande & Senzala: formação da família brasileira sob o regime de economia patriarcal. In: SANTIAGO, Silviano (coord.). **Intérpretes do Brasil**. Rio de Janeiro: Nova Aguilar, v. 2, p. 121-645, 2002.

HICKEY, Raymond. Contact and Language Shift. In: HICKEY, R. (ed.). **The Handbook of Language Contact**. Chichester: Wiley-Blackwell, p. 151-169, 2010.

INVERNO, Liliana. A transição de Angola para o português vernáculo: estudo morfossintático do sintagma nominal. In: CARVALHO, Ana M. (org.). **Português em Contato**. Madrid: Iberoamericana/Veuvert, p. 87-106, 2009.

LABOV, William. **Padrões sociolinguísticos**. 1 ed. São Paulo: Parábola, 2008 [1972].

LABOV, William. What can be learned about change in progress from synchrony descriptions. In: SANKOFF, David; CEDERGREN, Henrietta (Ed.). **Variation Omnibus**. Carbondale; Edmonton: Linguistic Research, p.177-199, 1981.

LABOV, William. **Principles of linguistic change**. Oxford; Cambridge: Blackwell, 1994.

LABOV, William. **Principles of linguistic change: social factors**. Oxford: Blackwell, 2001a.

LABOV, William. **Principles of linguistic change: cognitive and cultural factors**. Oxford: Wiley Blackwell, 2001b.

LABOV, William. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Org.). **Sociolinguistics: the Essential Readings**. Oxford: Blackwell, p. 235-250, 2003.

LABOV, William. **Principles of linguistic change; cognitive and cultural factors**. Oxford/Cambridge: Blackwell, v. 3, 2010.

LUCCHESI, Dante. **A variação na concordância de gênero em uma comunidade de fala afro-brasileira: novos elementos sobre a formação do português popular do Brasil**. 2000. Tese (Doutorado em Linguística) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

LUCCHESI, Dante. **Sistema, mudança e linguagem: um percurso na história da linguística moderna**. São Paulo: Parábola, 2004.

LUCCHESI, Dante. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto. In: RONCARATI, Cláudia; ABRAÇADO, Jussara (Org.). **Português brasileiro II: contato linguístico, heterogeneidade e história**. Niterói: EDUFF, 2008. p.366-390.

LUCCHESI, Dante. A concordância de gênero. In: LUCCHESI; Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). **O Português Afro-Brasileiro**. Salvador: Edufba, 2009a. p. 295-318.

LUCCHESI, Dante. História do Contato entre Línguas no Brasil. In: LUCCHESI; Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). **O Português Afro-Brasileiro**. Salvador: Edufba, 2009b. p. 41-73.

LUCCHESI, Dante. Conclusão. In: LUCCHESI; Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). **O Português Afro-Brasileiro**. Salvador: Edufba, 2009c. p. 513-546.

LUCCHESI, Dante. A diferenciação da língua portuguesa no Brasil e o contato entre língua. **Estudos de linguística galega**, Santiago de Compostela, v. 4, p. 45-65, 2012a.

LUCCHESI, Dante. A deriva secular na formação do português brasileiro: uma visão crítica. In: LOBO, Tânia; CARNEIRO, Zenaide; SOLEDADE, Juliana; ALMEIDA, Ariadne; RIBEIRO, Silvana (orgs.). **ROSAE: linguística histórica, história das línguas e outras histórias**. Salvador: EDUFBA, 2012b. p. 249-274.

LUCCHESI, Dante. O contato entre línguas e a origem do português brasileiro. In: GUGENBERGER, Eva; MONTEAGUDO, Henrique; REIDOUVAL, Gabriel. **Contacto de línguas, hibrididade, cambio: contextos, procesos e consecuencias**. Santiago de Compostela: Consello da Cultura Galega, 2013.

LUCCHESI, Dante. **Língua e Sociedade Partidas: a polarização sociolinguística do Brasil**. São Paulo: Contexto, 2015a.

LUCCHESI, Dante. O contato entre línguas na história sociolinguística do Brasil. In: VALENTE, André (Org.). **Unidade e Variação na Língua Portuguesa: suas representações**. São Paulo: Parábola, 2015b. p. 80-100.

LUCCHESI, Dante. Language in contact in Brazil and the genesis of creole languages. **Journal of Ibero-Romance Creoles**, v. 9, n. 1, p. 334-357, 2019a.

LUCCHESI, Dante. Por que a crioulização aconteceu no Caribe e não aconteceu no Brasil? Condicionamentos sócio-históricos. **Gragoatá**, Niterói, v.24, n. 48, p. 227-255, jan.-abr. 2019b.

LUCCHESI, Dante; BAXTER, Alan. A transmissão Linguística Irregular. In: LUCCHESI; Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). **O Português Afro-Brasileiro**. Salvador: Edufba, 2009. p. 101-124.

LUCCHESI; Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). **O Português Afro-Brasileiro**. Salvador: Edufba, 2009.

LUCCHESI, Dante; MACEDO, Alzira. A variação na concordância de gênero no português de contato do Alto Xingu. **Papia**, Brasília, v. 9, p. 20-36, 1997.

MAHER, Terezinha. Sendo índio em português... In: SIGNORINI, Inês (Org.). **Língua(gem) e identidade: elementos para uma discussão no campo aplicado**. Campinas: Mercado de Letras/Fapesp, 1998. p. 115-38.

MAIA, Marcus. A revitalização de línguas indígenas e seu desafio para a educação intercultural bilíngue. *Tellus*, ano 6, n. 11, p. 61-76, 2006

MALER, Bertil. **O Orto do Esposo**; glossário. Stokholm: Almqvist & Wiksell, 1964.

MATRAS, Yaron. Contact, convergence and typology. In: HICKEY, Raymond (Ed.). **The Handbook of Language Contact**. Chichester: Wiley-Blackwell, p. 66-85, 2010.

MCWHORTER, John. **Language Interrupted**. Oxford: Oxford University Press, 2007.

NARO, Anthony; SCHERRE, Marta. **Origens do português brasileiro**. São Paulo: Parábola, 2007.

PETTER, Margarida. As línguas africanas no Brasil. In: CARDOSO, Suzana; MOTA, Jacyra; MATTOS E SILVA, Rosa Virgínia (Org.). **Quinhentos Anos de História Linguística do Brasil**. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, p.117-141, 2006.

RODRIGUES, Aryon. Línguas indígenas: 500 anos de descobertas e perdas. **Delta**. São Paulo: v. 9, n. 1, p. 83-103, 1993.

RODRIGUES, A. D.; DIETRICH, W. On the linguistic relationship between mawé and tupi-guarani. **Diachronica**. Amsterdam: v. XIV, n. 2, p. 265-304, 1997.

ROMAINE, Suzanne. Contact and Language Death. In: HICKEY, Raymond (ed.). **The Handbook of Language Contact**. Chichester: Wiley-Blackwell, p. 320-339, 2010.

SANKOFF, David; TAGLIAMONTE, Sali; SMITH, Eric. **Goldvarb X: a variable rule application for Macintosh and Windows**. Toronto: Department of Linguistics, University of Toronto, 2005.

SCHERRE, Marta. **Reanálise da concordância nominal em português**. Rio de Janeiro: Universidade Federal do Rio de Janeiro, Tese de Doutorado em Linguística, 1988, 468 f.

SHERRE, Marta. Aspectos da concordância de número no português do Brasil. **Revista Internacional de Língua Portuguesa**. Lisboa: v. 12, p. 37-49, 1994.

SILVA, Raynice Geraldine Pereira da. **Estudo morfossintático da língua sateré-mawé**. Campinas: Universidade Estadual de Campinas, Tese de Doutorado em Linguística, 2010, 247 f.

SILVA NETO, Serafim da. **História da língua portuguesa**. 5 ed. Rio de Janeiro: Presença, 1988. TEIXEIRA, P. **Sateré-Mawé: retrato de um povo indígena**. Manaus: UFAM, 2005.

TRUDGILL, Peter. Contact and simplification: Historical baggage and directionality in linguistic change. **Linguistic Typology**, v. 5, n. 2-3, p. 371-374, 2001.

TRUDGILL, Peter. Sociolinguistic typology and complexification. In: SAMPSON, Geoffrey; GIL David; TRUDGILL, Peter (eds.). **Language Complexity as an Evolving Variable**. Oxford: Oxford University Press, p. 97-108, 2009.

TRUDGILL, Peter. Contact and Sociolinguistic Typology. In: HICKEY, Raymond (ed.). **The Handbook of Language Contact**. Chichester: Wiley-Blackwell, p. 299–319, 2010.

VEADO, Rosa Maria A. **Comportamento linguístico do dialeto rural**. Belo Horizonte: UFMG/PROED, 1982.

VOTRE, Sebastião Josué. Relevância da variável escolaridade. In: MOLLICA, Maria Cecília; BRAGA, Maria Luiza. **Introdução À Sociolinguística: o Tratamento da Variação**. 3. ed. São Paulo: Contexto, 2007.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. **Fundamentos empíricos para uma teoria da mudança lingüística**. 1 ed. São Paulo: Parábola, 2006[1968].

WINFORD, Donald. **An introduction to contact linguistic**. Oxford: Blackwell, 2003.