

**ABORDANDO A EMERGÊNCIA DAS PECULIARIDADES DO  
SUJEITO DO PORTUGUÊS BRASILEIRO A PARTIR DAS MARGENS  
DA CIÊNCIA LINGUÍSTICA**

**APPROACHING THE EMERGENCE OF THE PECULIARITIES OF THE  
BRAZILIAN PORTUGUESE SUBJECT FROM THE MARGINS  
OF LINGUISTIC SCIENCE**

ESMERALDA VAILATI NEGRÃO\*  
Universidade de São Paulo  
evnegrao@usp.br

EVANI VIOTTI  
Universidade de São Paulo  
viotti@usp.br

Este artigo tem como objetivo explorar algumas possibilidades de explicação das peculiaridades do português brasileiro a partir de certos temas que usualmente são tratados no âmbito de uma linguística que pode ser considerada periférica à linguística nuclear. Esses temas incluem: a caracterização das ecologias de comunicação como eminentemente multilíngues; o contato de línguas e variedades dialetais e idioletais como fator principal para qualquer mudança de línguas; a discussão sobre a classificação categórica de línguas feita em termos de grupos tipológicos claramente delimitados; a participação de línguas de outras famílias de línguas, para além da indo-europeia, nas ecologias em que se formaram as línguas coloniais, entre elas o português brasileiro. Por meio do exame da estrutura sintática dessa língua e de seu cotejo com construções de línguas africanas das famílias bantu e gbe, sugerimos que o português brasileiro deva ser considerado uma língua tipologicamente híbrida, na medida em que as peculiaridades de sua posição de sujeito a aproxima de famílias linguísticas que não a românica.

**Palavras-chave:** português brasileiro; línguas bantu; línguas gbe; multilinguismo; posição de sujeito

This article aims at exploring some possible explanations for the peculiarities of Brazilian Portuguese, from the perspective of some discussions which take place in the margins of nuclear linguistics. This perspective includes debates on themes such as: the characterization of ecologies of communicative practice as multilingual; the questioning of the categorical classifications of languages within clearly bounded typological groups; the participation of languages belonging to different linguistic families in the ecologies within which emerged the colonial languages, such as Brazilian Portuguese. By examining its

syntactic structure and contrasting it with constructions of African languages from the bantu and gbe families, we suggest that Brazilian Portuguese should be considered as a typologically hybrid language, inasmuch as the peculiarities of its subject position brings it closer to linguistic families other than that of Romance languages.

**Keywords:** Brazilian Portuguese; Bantu languages; Gbe languages; multilingualism; subject position

Recibido: 11 agosto 2020

Aceptado: 14 octubre 2020

## 1. INTRODUÇÃO

O objetivo deste artigo é o de discutir alguns dados relacionados à posição de sujeito do português brasileiro, a partir da hipótese de que sua gramática teria emergido no contexto colonial, caracterizado como uma ecologia multilíngue em que línguas de famílias diversas se encontravam em estreito contato.

Há alguns anos vimos investigando algumas construções do português brasileiro que exibem diferentes graus de impessoalização, desde sentenças existenciais e apresentacionais (Franchi *et al.* 1998; Negrão 1992; Viotti 1999), até construções que chamamos *absolutas*,<sup>1</sup> cuja sintaxe descrevemos e analisamos em Negrão e Viotti (2010). Essa análise sugere que essas sentenças decorrem de uma mudança de natureza semântica pela qual passa o português brasileiro, que demonstra uma preferência pela conceitualização de eventos que subjaz o sistema ergativo-absolutivo, maior do que aquela que é típica das demais línguas românicas, tipologicamente caracterizadas como línguas do sistema nominativo-acusativo. O sistema ergativo-absolutivo se caracteriza justamente por privilegiar a conceitualização e expressão do argumento temático do verbo, ou seja, aquele que recebe ou é afetado pela ação denotada pelo verbo, devendo marcar, em geral com um morfema de Caso ergativo, o argumento que desencadeia a ação verbal. Construções absolutas são aquelas que, tendo o sistema ergativo-absolutivo como sua base semântica, não só não expressam o argumento que se refere à fonte ou causa da ação verbal, como nem permitem sua conceitualização. São sentenças totalmente impessoais, que denotam eventos como se eles simplesmente acontecessem, sem que nada lhes tivesse servido como causa.

Em seguida, passamos a nos perguntar o que teria causado o aumento do uso dessa construção em português brasileiro. Em outros termos, o que teria feito o português brasileiro se afastar do padrão prototípico indo-europeu, mais especificamente românico, para começar a se aproximar de línguas que têm sido tipologicamente consideradas como membros de outras famílias? De um ponto de vista sócio-histórico, o português brasileiro, quando comparado às línguas românicas europeias, pode ser considerado uma língua recente, cuja gramática começou a emergir no período colonial, em um contexto de intenso multilinguismo. Além da língua

\* Esta pesquisadora recebe suporte do CNPq, com bolsa de produtividade em pesquisa, processos nº 307609/2014-6 e nº 309835/2018-6

<sup>1</sup> Essas sentenças correspondem ao membro monoargumental de um par de sentenças construídas com verbos que poderiam ser considerados verbos de alternância causativa. O que é digno de nota a respeito dessas sentenças no português brasileiro é que elas podem ser feitas a partir de um número de verbos muito maior do que aquele que caracteriza a classe de verbos de alternância causativa nas línguas românicas. Exemplos são: *Esse trem já perdeu; O concurso não anulou; Ele torceu pra o saque errar*, entre muitos outros.

trazida pelos exploradores e colonizadores portugueses em uma variedade de dialetos, falavam-se também aqui algumas outras línguas europeias, um expressivo número de línguas americanas nativas, e outro grande número de línguas trazidas pelo massivo número de africanos que foram forçadamente transportados para o território brasileiro durante o período do tráfico de escravizados.

Começamos, então, a explorar a hipótese de que o aumento de casos de construções absolutas que verificamos no português brasileiro poderia ter alguma relação com as línguas com as quais o português do século XV tenha tido contato no período colonial. Esse caminho foi percorrido em uma série de artigos (Negrão e Viotti 2008, 2011, 2014a, 2014b), em que investigamos a possibilidade de que a disseminação das sentenças absolutas no português brasileiro talvez pudesse ser explicada pela congruência de traços de construções com deslocamento à esquerda do português clássico, com traços de um tipo particular de construção do kimbundu, língua da família bantu falada no território que hoje corresponde a Angola, de onde veio um número substancial de africanos.

Desde o início dessa investigação, propusemos que o estudo da participação das línguas africanas na formação do português brasileiro fosse abordado a partir de uma perspectiva epistemológica que, à época, era pouco conhecida no Brasil (Negrão e Viotti 2011). Essa perspectiva é parte de um conjunto de estudos sobre a linguagem que se encontra às margens do núcleo da ciência linguística. Na próxima seção, vamos tratar dessa dicotomia epistemológica que separa o conhecimento que temos sobre a linguagem em um núcleo e uma periferia, e mostrar como uma abordagem vinda da marginália da linguística pode abrir novos caminhos de investigação das línguas humanas. Em seguida, tendo como pano de fundo algumas das questões levantadas no âmbito da marginália linguística, vamos voltar a tratar do impacto que línguas africanas tiveram na formação da sintaxe do português brasileiro, desta vez focalizando a posição de sujeito e a periferia esquerda da sentença, que codifica a estrutura informacional, e nos apoiando em construções de línguas das famílias gbe e bantu.

## 2. A MARGINÁLIA DOS ESTUDOS LINGUÍSTICOS

Em recente conferência proferida como parte da programação da *Abralin ao Vivo–Linguists Online*<sup>2</sup>, o africanista Felix Ameka nos lembrou de um antigo artigo de Brian D. Joseph, que sugere que, enquanto boa parte da atenção dos linguistas tem se voltado para alguns aspectos dos estudos da linguagem tidos como nucleares, a linguística deveria passar a levar em consideração aqueles aspectos da linguagem que têm sido considerados periféricos. Nas margens do núcleo da linguagem, há elementos de grande valor (Joseph 1997).

Não surpreende que essa ideia tenha sido retomada por um africanista. Afinal, as línguas africanas apresentam uma variedade de fenômenos que resistem a um enquadramento no arcabouço teórico da quase totalidade das teorias linguísticas correntes, dentre os quais destacam-se algumas interjeições e vocalizações (Ameka 1992, Ameka e Wilkins 2006), ideofones (Ameka 2001, Dingemanse 2017), logoforicidade (Ameka 2017), entre muitos outros.

Fenômenos como esses são exemplos de uma marginália que poderíamos caracterizar como aquela que não favorece generalizações, sempre tão caras a qualquer ciência, dificultando sua

<sup>2</sup> *Abralin ao Vivo – Linguists Online* foi uma série de conferências organizadas pela Associação Brasileira de Linguística (ABRALIN), com a colaboração de várias importantes associações de linguística do mundo, que ocorreu entre 04 de maio e 31 de julho de 2020.

integração a análises linguísticas mais amplas e abrangentes. Mas Joseph menciona ainda uma marginalia de caráter social que inclui alguns fenômenos que tangenciam a discussão que pretendemos fazer aqui. Entre eles, encontram-se aqueles relacionados a línguas emergentes (como, entre outras, aquelas que caracterizam a aquisição de língua estrangeira), e a línguas consideradas como simplificações de outras línguas (como muitos acreditam ser o caso de pidgins e crioulos).

Incluem-se, então, entre a marginalia de natureza social, alguns outros fenômenos linguísticos, como o multilinguismo e a mudança linguística causada pelo contato entre línguas que, a nosso ver, são relevantes para o entendimento da emergência do português brasileiro. Embora esses fenômenos venham sendo tratados por algumas vertentes da linguística nuclear, o arcabouço teórico-metodológico que norteia seu tratamento é ainda aquele que foi construído para dar conta dos fenômenos típicos de núcleo, e que tem por objetivo chegar a generalizações e abstrações, fazendo categorizações robustas com limites bem definidos. Como vamos ver adiante, esses dois fenômenos, entendidos a partir das margens da linguística, mostram-se mais desafiadores em relação a esse arcabouço teórico-metodológico.

De maneira geral, a linguística nuclear tem operado com a ideia de que, tanto no nível idioletal quanto no nível de línguas comunais, o monolingüismo é a norma, o multilingüismo é a exceção, e há uma clara demarcação separando uma língua de outra, ou um dialeto de outro. Essa ideia permeia as mais variadas áreas da linguística. De um lado, situam-se aquelas que, por exemplo, servem de referência a políticas educacionais favoráveis a que as crianças sejam conduzidas a preferir uma única língua (ou uma única variante de uma língua) –aquela que vai lhes garantir um maior acesso a boas oportunidades sócio-econômicas– abrindo mão de outras; ou ainda a tipologia, que, pela sua própria natureza, reduz a riqueza da variação que caracteriza o uso linguístico dos indivíduos de uma comunidade (Lüpke (2016). De outro lado, estão teorias linguísticas de base cognitiva, como, por exemplo, a chomskyana, que parte da hipótese de que a espécie humana se define por ser dotada geneticamente de uma capacidade cognitiva particular que vai orientar o processo de aquisição da linguagem de uma criança. Tipicamente, essa capacidade cognitiva é concebida para dar conta da aquisição de *uma* língua.<sup>3</sup> A partir dessa perspectiva, as categorizações opõem língua materna a segunda língua ou a língua estrangeira.

Abordar a questão a partir das margens da linguística implica rever essas ideologias. As separações categóricas entre línguas e dialetos, ou mesmo de línguas entre si e de dialetos entre si; as abstrações conducentes a uma reificação das línguas que não encontra correlatos diretos na fala; as noções de língua materna, segunda língua e língua estrangeira; e a própria caracterização da capacidade cognitiva para aquisição de língua ganham diferentes contornos quando outros contextos de prática linguística entram em jogo. Basta abrirmos nossas lentes para ver que há mais países multilíngues do que monolíngues, e até mesmo que falantes puramente monolíngues são raros (Aboh 2020a). Alguns lugares da África, por exemplo, caracterizam-se por ser ecologias multilíngues, em que cada língua serve a uma determinada função: uma é usada em rituais, outra é usada para o comércio, ainda uma outra é usada em família (Mufwene 2003; Cobbinah 2020). Mas há também ecologias em que a noção de que possa haver uma única língua dominando um certo uso linguístico sequer chega a existir, tal a fluidez que se verifica no uso de diferentes línguas e dialetos em uma mesma prática comunicativa. Lüpke (2016) chama

---

<sup>3</sup> Isso não quer dizer que a gramática gerativa entenda que não possa haver aquisição simultânea de mais de uma língua durante o período natural de aquisição de linguagem, aquele que vai do nascimento até, no máximo, a puberdade. Entretanto, essa teoria enfatiza a aquisição de *uma* primeira língua, considerando casos de bilingüismo ou de multilingüismo como exceções.

essas práticas de “falas multilíngues”. Elas evidenciam que, em contextos multilíngues, além de os falantes usarem diferentes línguas para diferentes fins, a permeabilidade que caracteriza os sistemas linguísticos –uma ideia que também é mais típica da periferia da linguística do que de seu núcleo– permite que as línguas se misturem dinamicamente em cada instância de uso (Lüpke 2016, Stenzel e Khoo 2016).

Tentar descrever a configuração linguística de contextos como esses em termos de língua materna ou língua nacional, e segunda língua ou língua estrangeira, revela-se uma limitação. Em alguns lugares da África, por exemplo, um determinado território terá uma língua patrimonial, que é a língua falada pelos antigos fundadores da comunidade, e que é a língua de seus descendentes, possíveis líderes contemporâneos daquele território. Outros grupos populacionais que tenham se juntado a essa comunidade como colonos podem até falar essa língua com alto grau de proficiência, mas não a consideram como sua. A identidade linguística desses colonos é associada à língua do lugar de onde eles vieram (Lüpke 2016, 2020). Em outros contextos, como é o caso da região do Alto Rio Negro, na Amazônia, há uma estreita ligação entre língua e identidade étnica<sup>4</sup>. Alguns povos, como, por exemplo, os da família tukano oriental, são exogâmicos, ou seja, casam-se com pessoas que falam línguas diferentes das suas. A filiação étnica e linguística é feita pela linha paterna: são tukanos e falantes de tukano aqueles que têm um pai tukano (Epps 2018). Mesmo que alguém seja fluente em tukano por ser essa a língua de sua mãe e não de seu pai, sua identidade linguística e étnica não é tukano. Fica claro, então, que, em contextos como esses, a pergunta a um falante sobre qual seria sua língua materna ou sua primeira língua não será pertinente (Cobbinah 2020). O impacto dessa constatação é de tal ordem que mesmo linguistas de orientação chomskyana, que assumem a hipótese da gramática universal como parte da dotação genética da espécie, têm sugerido que a capacidade humana para a linguagem deve ser reconsiderada de modo a que ela se defina como uma capacidade voltada para o multilinguismo, na medida em que falantes monolíngues ou monodialetais são raros, e que o que caracteriza o ser humano é justamente ser uma espécie capaz de lidar com traços linguísticos que pertencem a diferentes gramáticas, fazendo uso deles em contextos comunicativos específicos (Aboh 2015, 2020a, b).

No que diz respeito às análises tipológicas das línguas, é plausível pensar que boa parte dos falantes de quaisquer línguas não sabe dizer a que família sua(s) língua(s) pertence(m). No Brasil, muitas pessoas podem chegar a traçar a filiação do português brasileiro ao português europeu, mas apenas um número menor de falantes vai além disso. Em algumas ecologias que se caracterizam por um multilinguismo predominantemente indígena, como é o caso do Alto Rio Negro, por exemplo, os falantes podem não reconhecer ou acatar a divisão que os linguistas fazem das várias famílias linguísticas que coexistem na região, mas certamente separam as línguas indígenas das línguas dos brancos (Viotti 2017). Na África, a situação é diferente. Línguas indo-europeias faladas no continente, como, por exemplo, o português angolano, ou os crioulos e pidgins falados em diversos países africanos são percebidos pelos falantes como línguas africanas (Cobbinah 2020). Considerar as categorizações tipológicas das línguas por meio de uma perspectiva esboçada a partir da periferia nos faz ver, então, que existem outros modos de organização das línguas, baseados não em critérios genealógicos, mas, sim sócio-políticos. São as práticas sociais e as identidades étnicas e culturais que determinam o arranjo das várias línguas em uma determinada ecologia. Tudo isso pode impactar nosso entendimento sobre o contato de línguas, como vamos ver mais adiante. Sobretudo, a visão que os falantes

---

<sup>4</sup> Epps e Stenzel (2013) observam que essa associação entre língua e etnia tem se diluído mais recentemente em função da extinção de algumas línguas e do risco de extinção de outras.

têm do mapa linguístico que caracteriza a ecologia em que eles se situam é muito mais flexível do que a dos pesquisadores que trabalham no âmbito da linguística nuclear. Os falantes estão interessados em interagir e se comunicar eficientemente; para conseguir fazer isso, eles tornam bastante fluidos os limites entre os vários dialetos e as várias línguas (Cobbinah 2020).

A propósito da mudança linguística causada pelo contato entre línguas, por longo tempo a linguística histórica nuclear tem entendido a diversificação ou especiação das línguas como sendo fruto de pressões internas a um determinado sistema linguístico<sup>5</sup>. Essas pressões fariam com que as línguas se alterassem ao longo de suas histórias, assumindo características diferentes, vindo, em alguns casos, a dar origem a línguas diferentes<sup>6</sup>. O modelo que embasa essa visão é aquele proposto por Augusto Schleicher no século XIX, de acordo com o qual a diversificação de línguas considerada *normal* é aquela que ocorre monoparentalmente. Dentro dessa perspectiva, o contato entre línguas é visto apenas como um meio de explicar os desvios que resistem às explicações genéticas (Mufwene 2008: 30).

Herdeiro de uma ideologia de pureza que se forjou entre os séculos XVII e XIX, o modelo de Schleicher via como *anormal* qualquer mistura entre línguas, e também entre raças. O que dizer, então, das misturas de línguas que ocorrem nas situações de contato, e que impactam todas elas a ponto de modificá-las em diferentes graus, até aquele em que consideramos estar diante de uma nova língua? O que dizer das línguas crioulas, que além de se revelarem como resultados inquestionáveis do contato entre línguas (e raças), constituem casos extremos de hibridização linguística? Por que as variedades não crioulas que emergiram no contexto colonial, como é o caso do português brasileiro, são consideradas diversificações ‘regulares’ das línguas europeias, e as línguas crioulas, não (Mufwene 2008: 30)? Como explicar os fenômenos observados em ecologias multilíngues na África mencionadas acima, em que, em um mesmo enunciado, um falante faz uso de várias línguas diferentes?

Não há línguas puras, nem nunca houve. Embora o núcleo dos estudos da linguagem tenda a considerar que tenha havido uma “primeira” língua que teria emergido como um fenômeno que caracteriza a espécie, essa hipótese não é plausível. Existem evidências de que diferentes grupos de *homo sapiens* tenham coexistido simultaneamente<sup>7</sup>. Em qual desses grupos teria surgido essa “primeira” língua? E como essa língua teria se espalhado para os demais grupos, diversificando-se até chegar às línguas que conhecemos hoje? (Mufwene 2013: 53). Talvez um cenário mais adequado seja aquele em que se supõe que línguas diferentes tenham emergido mais ou menos concomitantemente em vários desses grupos e tenham se espalhado pelos territórios vizinhos, influenciando-se mutuamente e modificando-se à medida que seus falantes migravam de um lugar para outro. Línguas não são entidades autônomas: elas dependem das pessoas que as falam e acompanham essas pessoas em suas viagens; nesses movimentos, grupos populacionais encontram outros falantes de outras línguas, eventualmente juntam-se a eles formando novas comunidades e alterando as práticas linguísticas anteriores que caracterizavam ambos os grupos (Mufwene 2018: 388). A emergência das línguas românicas, por exemplo, é resultado da expansão de Roma. Soldados romanos, muitas vezes falantes de latim como

<sup>5</sup> O uso do termo “especiação” é um empréstimo da biologia para significar o processo evolutivo pelo qual novas línguas se formam.

<sup>6</sup> Sapir, por exemplo, pensa que as influências externas a uma língua podem afetar sua morfologia, mas apenas de maneira superficial. São as pressões internas a um sistema linguístico que devem ser responsáveis por quaisquer derivas da língua (Sapir 1949 [1921]: 150). Thomason e Kaufman (1988) e Thomason (2001), considerando que mudanças internas e externas têm natureza diferente, sugerem que só as primeiras podem ser consideradas a causa de especiação linguística. Ver Mufwene (2008) para uma crítica a essa abordagem.

<sup>7</sup> Ver Stringer (2012) citado em Mufwene (2013).

segunda língua ou língua estrangeira, carregaram consigo essa língua que ficou conhecida como latim vulgar por um vasto território, em que encontraram novos povos falantes de outras línguas. O contato entre esse latim vulgar e essa multiplicidade de outras línguas gerou mudanças no latim, até um ponto em que as línguas emergentes já tinham identidades mais ou menos caracterizadas como diferentes entre si. O mesmo pode-se dizer da emergência do inglês, que é consequência das invasões do território que hoje corresponde à Inglaterra por diversos povos germânicos e escandinavos. As línguas dos conquistadores entraram em contato com as línguas celtas faladas pelas populações que habitavam a região, já anteriormente expostas ao contato com a língua falada pelas legiões romanas que tinham dominado parte da Grã-Bretanha por aproximadamente cinco séculos. A África subsaariana, colonizada por populações bantu, passou por processo semelhante: o contato dessas populações com outros povos que viviam na área teve como resultado a diversificação das línguas bantu, e, concomitantemente, a perda gradual das línguas khoisan e pigmeu. Por esses exemplos, vê-se que, de maneira geral, a diversificação das línguas é antes fruto de movimentos populacionais e contatos de línguas, do que de mudanças internas ao sistema (Mufwene 2008: 31).

O contato de línguas visto a partir da perspectiva da linguística das margens implica a assunção de algumas ideias iniciais diferentes daquelas normalmente levadas em conta pela linguística nuclear. A primeira é a de que o tratamento das ecologias multilíngues deve ser feito de maneira diferenciada, de acordo com as especificidades de cada uma delas. Cada situação de contato se caracteriza por uma história socioeconômica particular que precisa ser cuidadosamente descrita e levada em consideração como um fator determinante para a explicação dos resultados do contato. A outra é a de que o contato de línguas envolve mais do que empréstimos de palavras e padrões de sentenças (Negrão e Viotti 2011, 2014b). De acordo com Mufwene (2008), a mudança linguística e a diversificação de línguas resultantes de situações de contato ocorrem por meio da combinação de dois fatores: (i) a seleção de alguns traços linguísticos, dentre várias alternativas em competição, que integram a gramática da chamada língua de superstrato, e (ii) a influência da(s) língua(s) de substrato na seleção de traços mencionadas em (i). Essa influência da(s) língua(s) de substrato está associada a condições de ordem demográfica e de ordem linguística. Do ponto de vista demográfico, quanto maior for o número de falantes da(s) língua(s) de substrato, mais forte deve vir a ser o impacto dessa(s) língua(s) sobre a língua emergente. Do ponto de vista linguístico, quanto maior for a congruência entre traços da(s) língua(s) de substrato e traços da língua de superstrato, maior vai ser a chance de esses traços serem selecionados como elementos da língua emergente.

No item a seguir, vamos apresentar e discutir algumas particularidades do português brasileiro que podem ser explicadas como fruto do contato do português com línguas africanas de diferentes famílias linguísticas, como a bantu e a gbe. Essa discussão evidencia o grande potencial descritivo e explicativo de uma visão de multilinguismo e contato de línguas mais aberta e dinâmica, como é aquela que procede de teorias que constituem a marginalia da teoria linguística, e que nos faz levar em consideração os seguintes fatores:

- O Brasil não é um país monolíngue, a despeito da hegemonia atual da variedade brasileira do português. Na época colonial, como já mencionado, o multilinguismo que caracterizava as práticas comunicativas diárias era ainda maior, na medida em que envolvia europeus falantes de diversas línguas, em especial o português; povos nativos que também falavam diferentes línguas, em especial as línguas da família tupi-guarani; e africanos que trouxeram consigo algumas línguas integrantes da grande variedade de línguas que caracteriza o continente africano. De particular interesse para este trabalho são línguas das famílias bantu e gbe. Dois documentos

históricos importantes dão suporte à hipótese de que línguas dessas famílias faziam parte da ecologia multilíngue do Brasil colônia: (i) *Arte da lingoa de Angola, oeferecida a Virgem Senhora N. do Rosario, Mãy, e Senhora dos mesmos Pretos*, primeira gramática da língua bantu kimbundu, escrita no Brasil, onde seu autor, o jesuíta Pedro Dias, nunca tendo estado em Angola, a aprendeu com os africanos que aqui viviam (Bonvini 2009; Rosa 2013); e (ii) *Obra nova de Lingoa geral de mina, traduzida, ao nosso Igdioma* (1741), escrita pelo português António da Costa Peixoto, que, à época, vivia no distrito de Vila Rica (atual Ouro Preto). Falada e documentada no Brasil, a língua geral de mina é uma língua africana semelhante a variedades de línguas gbe oriental, como o gungbe e o fongbe (Aboh 2015; Santos da Silva em preparação).

- Dentro desse quadro, é impossível desconsiderar o contato entre línguas como um fator determinante da emergência da variedade do português que começava a ser falada na colônia. Nesse sentido, o português brasileiro deve ser tomado como uma língua colonial, que nasceu em uma ecologia multilíngue caracterizada pelo intenso contato entre línguas. O estudo de línguas coloniais, sejam elas crioulas ou não, feito a partir de sua análise rigorosa, bem como a investigação das línguas que entraram em sua formação fazem ver como diferentes traços linguísticos, advindos de línguas tipologicamente distintas, se recombinaem e se reestruturam no processo de aquisição e construção das gramáticas híbridas que constituem os idioletos dos falantes (Aboh 2015). Além disso, as variedades emergentes não podem ser entendidas apenas como combinações particulares de traços das línguas que entraram em sua formação; esses traços são também eles modificados, para poder se conformar aos novos sistemas que estão emergindo. A partir dessa visão, cresce nosso conhecimento geral sobre como as línguas emergem, se modificam e evoluem (Mufwene 2001: 5).
- Do ponto de vista do contato com línguas africanas, é preciso considerar que os falantes dessas línguas vinham de uma ecologia em que o multilinguismo era (e é ainda hoje) a regra, e não a exceção, e que, em contextos multilíngues, o contato é generalizado.
- Grande parte do continente africano evidencia ainda hoje uma prática ancestral de comunicação multilíngue, em que a permeabilidade dos sistemas linguísticos é fortemente atestada. Falantes de línguas africanas parecem acostumados a trocar, em uma mesma prática comunicativa, uma língua pela outra, e a considerar como africana qualquer língua que eles usem em suas interações linguísticas, mesmo que essa língua tenha sido levada para o continente africano por conquistadores e colonizadores. É plausível supor, então, que os africanos que vieram ao Brasil no período colonial tivessem trazido consigo essa ideologia de flexibilização dos limites entre línguas e dialetos <sup>8</sup>.
- A ecologia multilíngue que se formou na colônia portuguesa na América se define por ter tido uma história particular, que não se caracterizou pela total separação de senhores e escravos, fossem eles indígenas ou africanos. Apesar de todo o vergonhoso abuso de poder dos senhores da terra sobre os escravizados e da violência desumana que caracterizou a escravidão no Brasil, as diferentes classes sociais mantinham interações comunicativas com relativa frequência. Embora essas interações fossem aquelas que caracterizam a comunicação entre falantes em processo de aquisição de língua estrangeira, a exposição desses falantes à língua portuguesa era razoavelmente constante.

---

<sup>8</sup> Na verdade, é pouco plausível que os próprios europeus que vieram para a colônia já tivessem, àquela época, uma noção nítida das fronteiras entre muitas línguas europeias. A visão de que a uma nação deve corresponder uma língua foi forjada a partir do século XVII.



Passemos, então, aos dados do português brasileiro que estamos considerando como possíveis focos do impacto que as línguas africanas tiveram português do século XV durante os séculos de colonização do Brasil.

### 3. A POSIÇÃO DE SUJEITO E A ESTRUTURA SENTENCIAL DO PORTUGUÊS BRASILEIRO: DO NÚCLEO DA LINGUÍSTICA EM DIREÇÃO A SUAS MARGENS

O interesse que a posição de sujeito do português brasileiro desperta nos estudiosos da linguagem vem de longa data. Do ponto de vista da linguística nuclear, a gramática gerativa tem, desde os anos 1980, procurado entender as peculiaridades do português brasileiro relativamente ao parâmetro do sujeito nulo, a partir das mudanças pelas quais ele teria passado, que teriam causado seu distanciamento em relação ao português europeu moderno e às demais línguas românicas<sup>9</sup>.

A partir do contraste relativo à distribuição e à interpretação de exemplos como (1) e (2) abaixo, Duarte (2000), entre outros, propõe que, enquanto o português europeu e outras línguas românicas, como o espanhol e o italiano, permitem o sujeito nulo com interpretação referencial, o português brasileiro apresenta uma forte tendência ao preenchimento da posição de sujeito

1. O nosso guarda-redes<sub>1</sub>, que horror! Na altura pro<sub>1</sub> já estava bastante magoado. pro<sub>1</sub> Tinha por título o pássaro de ferro. E pro sabe por que é que pro<sub>1</sub> se chamava pássaro de ferro?

(Português Europeu)

2. De repente ela<sub>1</sub> sabe que ela<sub>1</sub> quando criança ficava meio triste por isso.

(Português Brasileiro)

A sentença em (1) chega a levar os falantes de português brasileiro a ter problemas de interpretação, sendo difícil recuperar a relação entre os sujeitos vazios (pro) e o sintagma pleno a que correferem. A explicação corrente para esse fenômeno é a de que o português brasileiro, por ter passado por um processo de reestruturação de seu sistema pronominal, sofreu perdas nas distinções de pessoa e número da morfologia flexional do verbo (empobrecimento da flexão verbal), acarretando uma mudança paramétrica. Segundo essa mudança, o português brasileiro viria a deixar de ser uma língua de sujeito nulo (Galves 1993). Essa hipótese ganha reforço com a constatação da ocorrência de construções de deslocamento à esquerda, como (3), em que o constituinte deslocado é retomado por um pronome aberto na posição de sujeito. Esse tipo de construção não é encontrado em línguas de sujeito nulo<sup>10</sup>.

3. A Clarinha<sub>1</sub>, ela<sub>1</sub> cozinha que é uma maravilha.

<sup>9</sup> Para uma revisão cuidadosa e precisa dos fatos empíricos que sustentam essas análises, ver Pilati *et al.* (2017).

<sup>10</sup> Algumas das análises recentes propostas dentro do paradigma da gramática gerativa são: a hipótese da cisão da categoria pronominal no português brasileiro (Pilati e Naves 2013); construções tópico-sujeito (Kato 2015; Kato e Duarte 2018); cartografia de posições do sujeito no português brasileiro, além de uma possível flexibilização da grade temática de alguns verbos (Quarezemin e Cardinaletti 2017); mudança paramétrica nos padrões de concordância e Caso tratada no modelo de representação paramétrica (Avelar e Galves, 2016); entre outros.

Ainda de dentro da linguística nuclear, veio uma contestação para a hipótese de que o português brasileiro estaria em vias de deixar de ser uma língua de sujeito nulo, desta vez assentada sobre o fato de que há ainda muitos contextos em que o sujeito nulo é possível em português brasileiro, e que há uma assimetria de natureza sintático-semântica revelada pela interpretação que sujeitos nulos recebem, dependendo dos contextos sintáticos em que se encontram (Negrão 1997). Observem a sentença abaixo:

4. A Maria<sub>1</sub> disse que ela<sub>1/2</sub>/Ø<sub>1/\*2</sub> come chocolate quando ela<sub>1/2</sub>/Ø<sub>1/\*2</sub> está chateada.

Como se pode ver, tanto sujeitos nulos quanto pronomes plenos podem ocorrer em sentenças encaixadas. O pronome nulo na oração encaixada tem sua interpretação atrelada ao sujeito da sentença matriz. Diferentemente, quando os pronomes são plenos, eles podem tanto correferir com o sujeito da matriz, quanto podem ter uma interpretação dêitica dada pelo discurso.

Essa assimetria tem ainda consequências mais impactantes. Negrão (1997) observa que, quando o sujeito da matriz é um sintagma quantificado, o sujeito da sentença encaixada exibe uma distribuição complementar: apenas os sujeitos nulos poderão estar ligados ao sujeito da matriz; o sujeito pleno não poderá retomar a referência do sujeito da matriz, recebendo sempre uma interpretação dêitica que aponta para algum referente no discurso. É isso o que o exemplo a seguir captura:

5. Ninguém<sub>1</sub> acha que ele\*<sub>1/2</sub>/Ø<sub>1/\*2</sub> é inteligente.

Essa assimetria é claramente reafirmada pela diferença entre o comportamento de sujeitos nulos e plenos em sentenças encaixadas de construções com leituras *sloppy* e *estrita*. Observem o contraste entre (6) e (7):

6. Só o Maluf<sub>1</sub> acha que ec<sub>1</sub> vai ganhar as eleições.  
7. Só o Maluf<sub>1</sub> acha que ele<sub>1</sub> vai ganhar as eleições.

A interpretação da sentença (6) é *sloppy*: “o único x tal que x acha que x vai vencer as eleições é Maluf”, isto é, Maluf é o único candidato que pensa que ele mesmo vai vencer as eleições. Por outro lado, a interpretação de (7) é *estrita*: Maluf é a única pessoa que acha que o Maluf vai vencer as eleições. Nos contextos das sentenças (5) e (6), portanto, sujeitos vazios apresentam o comportamento de variáveis presas.

As evidências para essa distribuição complementar entre sujeitos nulos e plenos em sentenças encaixadas põem em xeque a hipótese de que o português brasileiro estaria deixando de ser uma língua de sujeito nulo. A assimetria é clara e mostra que o preenchimento da posição de sujeito no português brasileiro não pode ser explicado pelo empobrecimento da morfologia de sua flexão verbal. O que essa assimetria sugere é que o português brasileiro deva ser caracterizado como uma língua voltada para o discurso, em que sujeitos têm sua interpretação ligada a constituintes que ocupam alguma posição mais alta em que se codifica a estrutura informacional das sentenças (Negrão, 1999; Negrão e Viotti, 2000, Modesto, 2000a e 2000b).

A ideia de atrelar a posição de sujeito do português brasileiro à estrutura informacional das sentenças não é nova. A análise explorada pelo trabalho seminal de Eunice Pontes sobre tópicos

e sujeitos no português brasileiro aponta justamente para essa direção (Pontes 1986). De um ponto de vista empírico, as discussões sobre o sujeito do português brasileiro, a partir de variadas vertentes teóricas, têm tomado como ponto de partida os dados levantados por Pontes. Mas para além de se caracterizar como uma importante fonte de fatos que evidenciam as peculiaridades do português brasileiro moderno, esse trabalho faz uma análise inovadora dos constituintes das sentenças, tomando por base suas propriedades sintáticas, semânticas e informacionais, aliando-as aos padrões de ordem sentencial e concordância verbal. Enquanto os dados de Pontes entraram para o núcleo dos estudos sobre o português brasileiro, a análise propriamente dita parece ter sido deixada para as margens da linguística, e é justamente essa parte que queremos resgatar aqui: trata-se de uma análise que desmascara os problemas relativos à noção do sujeito do português brasileiro, tanto aqueles advindos do tratamento feito pelos gramáticos, quanto aqueles decorrentes das explicações elaboradas pelas teorias linguísticas. Mais ainda, a análise desvela as intuições de um grupo de falantes do português brasileiro sobre os padrões de estruturação e de concordância dessa língua, quando se defrontam com as sentenças peculiares colhidas por Pontes em sua pesquisa<sup>11</sup>.

Sentenças como (8), por exemplo, quando explicitadas pelos falantes, revelam a inadequação de análises que as tratam como instâncias de construções de deslocamento à esquerda ou de topicalização, na medida em que é arbitrário estabelecer a posição estrutural a partir da qual a ligação com o constituinte movido se dá:

8. Linguista a gente não pode conversar mais não.

Sem o contexto em que essa sentença foi enunciada, o linguista que procurasse explicar seu sentido poderia propor que ela fosse equivalente a uma sentença como (9), construída a partir do deslocamento do sintagma preposicional *com linguista* a partir de uma posição pós-verbal para o início da sentença, com a subsequente queda da preposição:

9. Com linguista a gente não pode conversar mais não.

Entretanto, levando em consideração o contexto e a interpretação do falante que proferiu a sentença, Pontes averiguou que a interpretação pretendida era: “com linguista por perto a gente não pode conversar porque ele anota o que se fala”.

Outro fator que norteou a análise de Pontes foi a resposta que ela obteve de um falante quando solicitado a fornecer uma explicação para a sentença (10), que ele havia acabado de produzir:

10. Esse rádio estragou o ponteiro.

O falante, que não era linguista nem da área de Letras, lhe respondeu que, nessa sentença, primeiro ele havia nomeado uma entidade –*esse rádio*– para avisar seu interlocutor que seria sobre essa entidade que ele iria falar, para em seguida poder falar sobre ela. Isso significa que o constituinte em posição pré-verbal é percebido por esse falante como aquele que codifica o

<sup>11</sup> Essas intuições foram obtidas através de (i) uma testagem feita com alunos de Letras (português e linguística) da Universidade Federal de Minas Gerais (UFMG); (ii) de uma pequena pesquisa com professores de português; e (iii) de conversas com seus colegas linguistas.

tópico a respeito do qual alguma coisa vai ser dita.

Essas experiências de cotejar as construções enunciadas com as intuições dos falantes que as produziram motivaram Pontes a concluir que construções como essas têm que ser analisadas dentro de uma teoria que leve em conta o contexto de sua produção, pois só assim será possível explicar a interpretação dada a elas pelo falante. E é só a partir dessa interpretação que o analista vai poder determinar com precisão quais são as estruturas sintáticas desses enunciados.

Algumas manifestações dos falantes que participaram da pesquisa apontaram, ainda, uma inconsistência no entendimento do que é o sujeito do português brasileiro. Como todos eles eram escolarizados, associar o sujeito ao agente do verbo foi uma conduta recorrente. Entretanto, quando os participantes foram questionados sobre as sentenças entre (11) e (14)<sup>12</sup>, eles consideraram que o constituinte pré-verbal com o qual o verbo concorda era indiscutivelmente o sujeito da sentença, independentemente de que, em nenhuma das sentenças, esse constituinte tinha uma semântica de agente. A pesquisa de Pontes mostra, assim, que a associação feita pelas teorias linguísticas entre a posição de sujeito e o papel temático de agente (ou causa) não é robusta em português brasileiro.

11. Essa casa bate bastante sol.

12. O João quebrou a perna.

13. Eu escorreguei o pé.

14. O João operou o estômago.

Pontes conclui então que a identificação do sujeito no português brasileiro se faz exclusivamente por ele ocupar uma posição pré-verbal na sentença e pela concordância marcada no verbo. Os papéis semânticos associados a esse constituinte –ser agente, ser tema, ser experienciador, ser locativo, ser beneficiário– são totalmente secundários (Pontes 1986: 25). Mais do que relações semânticas, as sentenças do português brasileiros se estruturam a partir de um conteúdo informacional, em um padrão tópico-comentário (Pontes 1986: 175): em posição pré-verbal vem o tópico, a informação velha, aquilo a respeito do que o falante vai tratar no comentário; e, em posição pós-verbal, vem aquilo que é informação nova.

Todas as línguas têm estratégias próprias para ligar o conteúdo proposicional de suas sentenças ao discurso, exibindo diferentes maneiras de codificar sua estrutura informacional: por alterações na ordem dos constituintes sintáticos, por prosódia, ou por construções específicas, como clivadas, por exemplo (Lambrecht 1994). O que parece ser uma característica marcante no português brasileiro é que, para além dessas estratégias, a própria posição de sujeito é prioritariamente percebida e utilizada como um locus informacional, mais do que como uma posição temática. Ou seja, enquanto nas línguas românicas em geral –no português europeu em particular– a posição de sujeito de uma sentença ativa é uma posição que expressa, de forma balanceada, um papel temático (fundamentalmente agente, causa, fonte da ação) e um conteúdo informacional (tópico), no português brasileiro essa posição tem como função principal a veiculação de uma certa estrutura informacional, abrindo-se, assim, para constituintes com papéis semânticos os mais diversos. O que pretendemos fazer aqui é mostrar que o contato com línguas africanas pode ter sido um fator que contribuiu sobremaneira para essa caracterização da posição de sujeito do português brasileiro.

<sup>12</sup> Essas sentenças não estavam entre os enunciados produzidos por eles, nem se encaixavam nas definições que eles próprios tinham dado para a noção de sujeito.

#### 4. NAS MARGENS DA LINGUÍSTICA. PARTE 1: AS CONSTRUÇÕES DE INVERSÃO DO SUJEITO E OS PADRÕES DE CONCORDÂNCIA NAS LÍNGUAS BANTU

Vamos, agora, dar início a uma aproximação entre as peculiaridades dos sujeitos do português brasileiro e algumas características das línguas bantu, que nos parece terem tido papel relevante na formação da gramática do português brasileiro.

As línguas bantu são faladas numa extensa área do continente africano, desde o sudeste da Somália a leste, e a República dos Camarões a oeste, até o extremo sul da África do Sul. Muitos africanos trazidos para o Brasil eram falantes de línguas pertencentes a essa família linguística. Apesar do grande número de línguas que constituem essa família e da larga extensão de território em que são faladas, as línguas bantu caracterizam-se por ser SVO, embora possam exibir diferentes ordens parcialmente motivadas pelo conteúdo informacional associado aos constituintes sentenciais. Além disso, elas compartilham certas uniformidades morfossintáticas que as caracterizam como uma única família linguística. Dentre essas uniformidades, duas são de particular interesse para a argumentação que estamos buscando fazer aqui: o sistema de concordância e as construções de inversão.

No que diz respeito à concordância, os substantivos dessas línguas são classificados em subclasses de acordo com um rico sistema de prefixos afixados a raízes nominais, conhecido como *classes nominais*. As classes nominais desempenham um papel importante nos padrões de concordância dessas línguas, uma vez que aos radicais verbais são prefixados afixos que expressam relações de concordância com a pessoa ou a classe nominal do constituinte em posição de sujeito. Esses afixos no verbo são chamados *marca de sujeito* (SM) e são obrigatórios. Outros afixos expressando concordância em pessoa e classe nominal com o constituinte em posição de objeto (OM) também se juntam ao verbo. Esses, no entanto, não são obrigatórios, e estão sujeitos a condições de licenciamento com variação entre as línguas. Entre a marca de sujeito e a marca de objeto no verbo aparecem prefixos marcadores das categorias gramaticais de tempo, aspecto e modo (TAM). As classes nominais distribuem-se em pares, em que um membro expressa singular e o outro plural. A sentença do swahili abaixo exemplifica a ordem canônica das sentenças, SVO, e o padrão de concordância comum às línguas bantu:

15. Wa-toto wa-le wa- na- ya- on - a ma-chungwa ha-ya<sup>13</sup>  
 2-criança 2- aquela SM2- PRES-OM6-ver-FV 6-laranja esta-6  
 ‘Aqueles meninas veem esta laranja’

O substantivo *menina* é marcado com o prefixo de classe nominal 2, plural, com o qual concordam tanto o demonstrativo *aquela*, marcado com classe 2, como o prefixo de marca de sujeito (SM2) no verbo. Com o substantivo em posição pós-verbal, marcado pelo prefixo de classe 6, concordam tanto o demonstrativo *esta*, também marcado com classe 6, como o verbo, marcado com o prefixo de objeto classe 6 (OM6).

No que diz respeito às construções de inversão de sujeito das línguas bantu, elas parecem estar diretamente relacionadas à noção de sujeito dessas línguas, que envolve, entre outros aspectos, a dificuldade em distinguir a função de tópico ou sujeito para o constituinte em posição pré-verbal. Essa sugestão vem de um grande estudo tipológico conduzido por Martin e

<sup>13</sup> Este dado foi retirado de Krifka (1995), com glosas adaptadas e traduzidas.

van der Wal (2014), a partir de dados de 46 línguas espalhadas por diversos países<sup>14</sup>, todas pertencentes à família bantu. Os autores corroboram a afirmação já feita em Baker (2008), de que, nessas línguas não há requerimento de que os sintagmas nominais sejam marcados com Caso<sup>15</sup>.

São sete os tipos de construção de inversão do sujeito: a inversão locativa formal; a inversão locativa semântica; a inversão do instrumento; a inversão do paciente; a inversão do complemento; a inversão com concordância default; e a concordância com o sujeito lógico invertido. As sentenças em (16) e (17) são um exemplo de inversão locativa formal em otjiherero, uma língua falada na Namíbia, em Botswana e no sudeste de Angola, próximo à fronteira com esses países:

16. È-rúngá      r-á      hití      m-ón-djúwó  
 5-ladrão    SM5-PST    entrar      18-9-casa  
 ‘O ladrão entrou na casa’
17. M-ón-djúwó    mw-á      hití      è-rúngá  
 18-9-casa      SM18-PST    entrar      5-ladrão  
 ‘Na casa entrou o ladrão’

A sentença (17) compartilha, com a inversão locativa formal de outras línguas bantu e com os outros tipos de inversões nessas línguas, as seguintes propriedades: (i) o sujeito lógico segue o verbo e não pode ser omitido; (ii) há uma relação de concordância entre o prefixo de classe nominal do constituinte pré-verbal, no caso a marca locativa 18 prefixada ao nome, e a marca de sujeito prefixada à raiz verbal, no exemplo (SM18); (iii) do ponto de vista informacional, o constituinte pré-verbal serve como tópico, o sujeito lógico pós-verbal introduz foco do tipo informação nova, ou foco apresentacional ou tético; (iv) a introdução do prefixo de marcação de objeto (OM) gera agramaticalidade; (v) há uma adjacência estrita entre o verbo e o sujeito lógico em posição pós-verbal, marcada por uma variedade de estratégias de ordem fonológica ou morfológica.

Por outro lado, as línguas bantu divergem quanto ao tipo de inversão de sujeito nelas encontrado. A inversão locativa formal, como a exemplificada acima, é a construção de inversão de sujeito mais comum entre as línguas estudadas. Se a ela agregarmos a inversão locativa semântica<sup>16</sup>, chegamos à conclusão de que a inversão locativa é caracterizadora dessas línguas. A sentença (18), da língua zulu, é um exemplo de inversão locativa semântica.

<sup>14</sup> Infelizmente, as línguas bantu das zonas F e H não foram integradas ao estudo. A zona H abarca a maioria das línguas de Angola, aí incluído o kimbundu, que, como visto, é uma das línguas bantu que foi falada no Brasil colônia. Estudos sobre a sintaxe do kimbundu e de outras línguas de Angola são escassos.

<sup>15</sup> O trabalho de Baker discute a comparação entre línguas da família indo-europeia e línguas bantu, mostrando que recorrentes problemas encontrados nessa comparação seriam resolvidos se dissociássemos Caso e concordância. As línguas bantu têm um rico sistema de concordância, que não envolve a marcação morfológica de Caso para a distinção entre as funções gramaticais dos constituintes da sentença.

<sup>16</sup> A inversão locativa semântica parece estar em distribuição complementar com a inversão locativa formal nas línguas para as quais estudos sobre essas construções são encontrados.

18. Le-zi      zi-ndlu      zi-hlala      aba-ntu      aba-dala<sup>17</sup>  
 DEM-10   10.casas   SM10-viver   2-gente      2-velho  
 ‘Essas casas vivem pessoas velhas’

Nessa sentença, a marca de classe locativa não aparece, o que leva a concordância no verbo a ser marcada pelo prefixo de classe do constituinte pré-verbal (SM10). Mesmo assim, esse constituinte é semanticamente um locativo.

Vejamos agora outros tipos de construções de inversão. As sentenças (19), da língua siswati e (20), da língua kirundi, são classificadas como casos de inversão do instrumento. Já as sentenças (21), da língua swahili e (22) da língua kirundi, são casos de inversão do paciente:

19. Le-moto      le-na      i-to-homb-a      tsine      kuphela<sup>18</sup>  
 DEM-9.carro   DEM-9   SM9-FUT-ir-FV   pron1pp      só  
 ‘Este carro viaja só nós’
20. I-karamu   y-   andikish-a      John<sup>19</sup>  
 9.caneta   SM9-escrever-FV   1.John  
 ‘A caneta escreve João’
21. Wimbo      u-ta-imb-a      wa-tu      mia<sup>20</sup>  
 11.música   11-FUT-canta-FV   2-pessoa      cem  
 ‘A música vai cantar cem pessoas’
22. Ama-tá      y-á-nyôye      abana<sup>21</sup>  
 6-leite      SM6-PST-beber.PRF      crianças  
 ‘O leite bebeu as crianças (não os pais)’

Em todas essas sentenças, o sujeito lógico é realizado em posição pós-verbal, podendo até expressar foco: a sentença (19) expressa um enunciado tético, e na sentença (22) o constituinte pós-verbal tem foco contrastivo. Por sua vez, os constituintes pré-verbais, desempenhando diferentes os papéis semânticos mencionados acima (instrumento e paciente) desencadeiam concordância de classe com o marcador de sujeito.

Esse tipo de fenômeno tipicamente bantu não parece diferir daquele atestado por Pontes em sua pesquisa sobre o sujeito no português brasileiro. Sentenças como aquelas coletadas por ela, elencadas entre (11) e (14) acima, consideradas como particulares ao português brasileiro, exibem características semelhantes às das sentenças de línguas bantu arroladas entre (19) e (22). Tanto quanto em bantu, o português brasileiro pode ter, em posição pré-verbal, constituintes com semântica de locativo, experienciador, possessivo, beneficiário/afetado, entre outros, e manifestar concordância com o verbo. O que importa, para as línguas bantu e para o português

<sup>17</sup> Dado retirado de Buell (2007:111) citado por Marten e van der Wal (2014), adaptado por nós.

<sup>18</sup> Dado retirado de Thwala (2006b:213) citado por Marten e van der Wal (2014), adaptado por nós.

<sup>19</sup> Dado de Ferdinand Mberamihigo (p.c.), citado por Marten e van der Wal (2014) e adaptado por nós.

<sup>20</sup> Dado retirado de Whiteley e Mganga (1969:113), e Russell (1985:477), citados por Marten e van der Wal (2014), adaptado por nós.

<sup>21</sup> Dado retirado de Ndayiragije (1999:400) citado por Marten e van der Wal (2014), adaptado por nós.

brasileiro parece ser a estrutura da informação. A posição pré-verbal de suas sentenças, que gera a concordância com o verbo, é uma posição em que se codifica o tópico da sentença, não uma posição caracterizada pela atribuição de um papel temático específico para o sujeito.

Além desses casos de proximidade entre a variedade brasileira do português e de línguas da família bantu, há outros igualmente indicadores de que essas línguas tiveram um forte impacto na constituição da gramática do português brasileiro. Trata-se de um outro fenômeno de inversão, que Marten e van der Wal (2014) chamam *inversão de concordância default*, e que compreende, entre outras, as construções de foco apresentacional. Novamente, estamos diante de um fenômeno que diz respeito à estrutura informacional das sentenças, e que coloca as relações entre a semântica do verbo e seus argumentos em uma posição secundária.

Na língua bantu setswana, dentre os prefixos marcadores de classes nominais nos nomes, há três prefixos locativos: *fa* (classe 16), *ko* (classe 17) e *mo* (classe 18). Segundo o padrão bantu de inversão locativa formal, os sintagmas nominais locativos marcados com esses prefixos podem aparecer tanto em posição pós-verbal, quanto em posição pré-verbal. E embora haja muita variação, em muitas línguas bantu a concordância entre as diferentes marcas de locativo do constituinte pré-verbal e as marcas que aparecem no verbo é consistente. Entretanto, em setswana, nas construções de inversão locativa formal, o prefixo de marca de concordância de sujeito no verbo é invariavelmente o da classe nominal 17, independentemente da marca de classe nominal locativa que marca o constituinte pré-verbal (considerado o sujeito), como mostram as sentenças abaixo<sup>22</sup>. Para além disso, nessas sentenças em que a concordância de marca de sujeito no verbo é invariavelmente o prefixo da classe 17, o sujeito lógico em posição pós-verbal tem foco apresentacional (Demuth e Mmusi 1997).

- |     |                                 |                |              |
|-----|---------------------------------|----------------|--------------|
| 23. | Fá-se-tlharé-ng                 | gó-émé         | ba-símané    |
|     | 16-7-árvore-LOC                 | SM17-estar/PRF | 2-meninos    |
|     | ‘Sob a árvore está os meninos’  |                |              |
| 24. | Kó-Maúng                        | gó-tlá-ya      | roná maríga  |
|     | 17-Maung                        | SM17-FUT-go    | 1pDM inverno |
|     | ‘Para Maung irá nós no inverno’ |                |              |
| 25. | Mó-le-fátshé-ng                 | gó-fúla        | di-kgomo     |
|     | 18-5-campo                      | SM17- pastar   | 10-gado      |
|     | ‘No campo pasta o gado’         |                |              |

Por fim, em setswana, o sintagma locativo em posição de sujeito pode ser omitido, como na construção abaixo:

- |     |                |          |
|-----|----------------|----------|
| 26. | Gó-fúla        | di kgomó |
|     | SM17-pastar    | 10-gado  |
|     | ‘Pasta o gado’ |          |

<sup>22</sup> Os exemplos são de Demuth e Mmusi (1997), citados em Marten e van der Wal (2014).



Nesses casos, a interpretação locativa se perde, e a marca 17 de concordância do sujeito no verbo passa a funcionar como uma concordância default<sup>23</sup>.

Uma vez mais, estamos diante de construções possíveis na família bantu que são reminiscentes de algumas construções impessoais do português brasileiro, como as seguintes<sup>24</sup>:

27. Existe muitos outros meios de transporte que não são explorados. (PA - 283).
28. Aí então começou a aparecer os vestidos feitos, mas muito pouco e pouca gente usava vestidos. (SP, 396).
29. Eu vou de moto, eu choro. Sai lágrimas. Bom é que estou andando sem óculos (SP - 343).
30. Foi uma sorte que foi localizado [o estouro das bombas] [...] Morreu cem mil de uma vez...(SP, 343).
31. Dá umas nanicas enormes na minha chácara.
32. Molhou tudo os quartos em casa na tempestade de ontem.

Essas construções são todas de foco apresentacional e não expressam o locativo, mantendo a posição de sujeito vazia, justamente como a construção do setswana em (26). Além disso, todas essas sentenças aceitam um constituinte locativo em posição pré-verbal como nos exemplos entre (23) e (25) do setswana. Também de forma análoga ao setswana, o verbo não concorda com o sujeito lógico das sentenças, que aparece em posição pós-verbal. É isso o que mostram os exemplos entre (33) e (38):

33. No país existe muitos outros meios de transporte que não são explorados.
34. Aí então em São Paulo começou a aparecer os vestidos feitos.
35. Dos meus olhos sai lágrimas.
36. No acidente morreu cem mil de uma vez.
37. Na minha chácara dá umas nanicas enormes.
38. Em casa molhou tudo os quartos na tempestade de ontem.

Para sumarizar esta nossa primeira incursão pelas margens da linguística na tentativa de entender fenômenos que expliquem a emergência de certas peculiaridades do português brasileiro, podemos dizer que o estudo da tipologia das construções de inversão do sujeito nas línguas bantu nos permitiu ver que a ordem dos constituintes das sentenças dessas línguas é fortemente regida pela natureza do conteúdo informacional que expressam, e que essa variação de ordem está atrelada a padrões de concordância que refletem a morfologia exuberante que caracterizam essas línguas. Embora o português brasileiro não exiba uma morfologia tão complexa quanto a das línguas bantu, parece-nos que ele compartilha com elas a preferência

---

<sup>23</sup> Marten e van der Wal (2014) sugerem que essa marca de concordância de sujeito nas construções de inversão com concordância default são historicamente marcadores de classe locativa que se gramaticalizaram como marcas default de concordância de sujeito.

<sup>24</sup> Esses exemplos foram discutidos em Franchi, Negrão e Viotti (1998). Aqueles entre (28) e (31) foram retirados do corpus NURC; os de número (32) e (33) são exemplos de ocorrências anotadas por Carlos Franchi.

pela codificação do conteúdo informacional. Essa preferência tem como consequência a possibilidade de que constituintes com semânticas diversas possam ocupar a posição de sujeito, e de que o verbo concorde com esses constituintes, e não com o sujeito lógico da sentença. Além disso, há uma semelhança evidente entre certas construções impessoais do português brasileiro, e construções de inversão com concordância default possíveis em muitas das línguas bantu.

## 5. NAS MARGENS DA LINGUÍSTICA. PARTE 2: A ESTRUTURA INFORMACIONAL E O COMPORTAMENTO PRONOMINAL DAS LÍNGUAS GBE.

Outras línguas africanas que atestadamente foram faladas no Brasil colônia pertencem ao grupo gbe, um subgrupo das línguas kwa, faladas no oeste da África, na região que hoje corresponde aproximadamente a Ghana, ao Togo e ao Benin. Muitos africanos foram trazidos dessa região para o Brasil, em particular aqueles que vieram para Minas Gerais para o trabalho de mineração durante o Ciclo do Ouro.

Como as línguas bantu, as línguas gbe são SVO. Entretanto, elas apresentam propriedades muito diferentes das que caracterizam as línguas bantu: (i) elas não admitem sujeito nulo, de modo que o preenchimento da posição do sujeito é obrigatório; e (ii) no que concerne à morfologia de concordância verbal, o verbo não carrega marcas de pessoa, gênero ou número, como evidenciado pelos exemplos de gungbe abaixo (Aboh 2005:154):

39. é            h ò n  
       3sg.    sair  
       ‘Ele/ela saiu’
40. yé        h ò n  
       3pl.    sair  
       ‘Eles/elas saíram’

Além disso, uma grande diferença entre as línguas bantu e as línguas gbe está no fato de que as duas famílias linguísticas expressam o conteúdo informacional de seus constituintes sentenciais de forma bastante diferente (Aboh 2007). Tópico e foco são expressos em dois domínios periféricos distintos: enquanto as línguas bantu usam o domínio periférico de VP para marcar constituintes topicalizados e focalizados e, em alguns casos, a periferia esquerda das sentenças, as línguas kwa (nelas contidas as línguas gbe) usam a periferia esquerda das sentenças, domínio periférico de CP, “de modo transparente”. Aboh usa essa transparência das línguas gbe como evidência para cada uma das categorias funcionais propostas pelo modelo cartográfico (Rizzi 1997), uma vez que essas línguas apresentam marcadores abertos que seguem os constituintes topicalizados e focalizados, sendo assim candidatos a ocupar a posição de núcleo das projeções funcionais TopP e FocP, respectivamente<sup>25</sup>. As sentenças da língua

<sup>25</sup> Para Ameka (2009: 145), dado que as línguas kwa apresentam posições estruturais específicas para constituintes exprimindo tópico e foco, elas podem ser consideradas “línguas de configuração discursiva”.

gungbe abaixo exemplificam a distribuição estrutural dos constituintes que exprimem essas relações discursivas (Aboh 2004: 309 e 242).

41. vi ló yà é yì wéxòme  
 criança Spf<sub>[+def]</sub> Top 3W-sg ir-Perf escola  
 ‘Criança, ele/ela foi à escola’
42. Rɛmi<sub>i</sub> wɛ t<sub>i</sub> gbá xwè  
 Remi Foc construir-Perf casa  
 ‘REMI construiu a casa’

Vejam que na sentença (41), o constituinte *criança* é realizado na periferia à esquerda da sentença seguido do marcador de tópico *yà*, e é retomado por um pronome resumptivo na posição de sujeito (analisado por Aboh como um pronome fraco). Trata-se, assim, de uma construção análoga a construções muito comuns do português brasileiro moderno – as construções de deslocamento à esquerda– como exemplificadas em (3).

Uma importante característica das línguas gbe que consideramos relevante para o entendimento da gramática do sujeito do português brasileiro é que o sistema pronominal dessas línguas pode exibir logoforicidade. Em ewe, outra língua gbe, o pronome logofórico é *ye*: ele ocorre como “um pronominal preso que se refere àquele cuja fala e pensamentos está sendo reportado” (Essegbey 1999:48).

Pronomes logofóricos são tradicionalmente caracterizados como tendo que ocorrer sob o escopo de predicados de atitude como *acreditar* e *dizer*, sendo obrigatoriamente ligados a um antecedente que é o portador dessa atitude. Além do logofórico *ye*, ewe possui um pronome de terceira pessoa *é*, que exibe as propriedades interpretativas normais de um pronome (Pearson, 2015: 78), como verificado abaixo.

43. Kofi<sub>1</sub> be yè<sub>1/\*2</sub> dzo  
 K. dizer LOG sair  
 ‘Kofi disse que saiu’
44. \*Yè dzo  
 LOG sair
45. Kofi<sub>1</sub> be e<sub>1/2</sub> dzo  
 K. dizer 3SG sair  
 ‘Kofi disse que ele saiu’
46. E dzo  
 3SG sai  
 ‘Ele saiu’

Em (43), o pronome logofórico na sentença subordinada obrigatoriamente toma como seu antecedente o sujeito da matriz, sendo ligado a ele. Se cotejarmos esse comportamento dos

pronomes do ewe com os exemplos do português brasileiro entre (4) e (7), vemos que ele é bastante semelhante à distribuição complementar observada entre o sujeito nulo e o sujeito aberto do português brasileiro moderno: o pronome nulo, em português brasileiro, parece que prefere se comportar como uma variável presa, cuja interpretação depende de constituintes que estão em posição mais alta na estrutura, tanto quanto os logofóricos do ewe, enquanto o pronome pleno mantém as características típicas de pronomes.

Mais uma vez, quando nos debruçamos sobre a análise de alguns aspectos da gramática das famílias das línguas africanas que foram trazidas para o Brasil durante o período colonial, encontramos uma série de características que se assemelham a algumas das peculiaridades do português brasileiro fundamentais para a argumentação de que sua gramática difere da do português europeu. A proximidade entre línguas gbe e o português brasileiro, no que diz respeito ao comportamento do sujeito, se observa em dois tipos de fenômeno: construções com deslocamento à esquerda, em que o constituinte deslocado é retomado na posição de sujeito por um pronome resumptivo, ressaltando a relevância da estrutura informacional para ambas as línguas; e a assimetria do sistema pronominal, em que alguns pronomes se comportam como variáveis ligadas (ou pronomes logofóricos), enquanto outros seguem o comportamento padrão dos pronomes.

## 6. CONSIDERAÇÕES FINAIS

A historiografia linguística e as teorias linguísticas contemporâneas têm apontado as inconsistências das primeiras gramáticas de línguas indígenas, africanas, asiáticas e do Pacífico quando, em sua metodologia e em suas descrições e explicações, seguem o modelo gramatical greco-latino (cf. Altman 2011; Zwartjes 2011)<sup>26</sup>. Esse é o caso da famosa *A Arte da língua mais falada no Brasil*, de José de Anchieta, ou d' *A Arte da Língua de Angola*, de Pedro Dias, mencionada acima. Entretanto, no incessante trabalho de descrição e análise de línguas feito no ocidente, muitos de nós continuamos a analisar as mais diversas línguas, pertencentes às mais variadas famílias linguísticas, valendo-nos do mesmo instrumental teórico-metodológico desenvolvido para a descrição de línguas indo-europeias<sup>27</sup>. O alerta mencionado na nota de rodapé 15 acima, feito por Baker a respeito do entendimento da morfologia das línguas africanas, aponta justamente para isso: analisar línguas da família bantu a partir da relação entre Caso e concordância usada para explicar as línguas indo-europeias acaba por distorcer nosso entendimento a respeito da gramática daquelas línguas. É importante descrever e analisar cada língua, ou cada família de línguas, a partir de unidades, categorias e relações gramaticais que lhes sejam próprias, e que não necessariamente correspondam ou tenham correlatos em outras línguas.

Por que estamos trazendo esse tipo de discussão para um trabalho sobre o português brasileiro, que, como língua indo-europeia que é, poderia, em princípio, ser descrita pelo mesmo aparato teórico usado para a descrição das demais línguas daquela família? É porque optamos por adotar uma perspectiva que vem da marginália da linguística, segundo a qual novas línguas emergem por causa do contato de línguas que ocorre quando populações de diferentes origens, histórias, culturas e línguas se encontram e convivem em uma ecologia que acaba se

<sup>26</sup> Agradecemos a Olga Ferreira Coelho Sansone pelas referências.

<sup>27</sup> Sobre uma discussão sobre o uso desse instrumental para a descrição de línguas indígenas, ver Negrão e Viotti (2014).

caracterizando por ser multilíngue. O português brasileiro é uma língua recente, que nasceu e viveu a primeira fase de seu desenvolvimento durante os séculos coloniais, em que pessoas falantes de línguas europeias, notadamente as diferentes variedades dialetais do português que foram trazidas para cá pelos exploradores e colonizadores europeus, conviveram com um número expressivo de pessoas falantes de línguas nativas brasileiras e de uma vasta gama de línguas africanas, especialmente as que pertenciam às famílias bantu e gbe<sup>28</sup>. Da perspectiva de uma linguística das margens, então, a explicação para as peculiaridades do português brasileiro deve necessariamente passar pela discussão do possível impacto que essas outras línguas teriam tido na constituição da nova gramática que emergia a partir da interação entre os falantes do Brasil colonial.

De maneira geral, quando as descrições e análises levam o contato linguístico em consideração, elas tomam como dados de interesse aqueles que podem ser explicados como empréstimos lexicais ou estruturais, sem que se façam discussões mais detalhadas sobre o arcabouço da gramática de onde vieram esses empréstimos. Crucialmente, do ponto de vista de uma linguística das margens, são esses aspectos que acabam por revelar o entendimento de partes importantes das gramáticas das línguas e variedades emergentes, que se configuram como algo inteiramente novo, não encontrável *ipsis verbis* em nenhuma das línguas em contato a partir das quais as novas variedades emergiram. Mesmo assim, esses novos fenômenos se relacionam com traços da gramática das línguas que lhes deram origem por um certo tipo de *semelhança de família* (Mufwene 2008).

Em trabalhos anteriores (Negrão e Viotti 2008, 2010, 2011, 2014a; Petter *et al.* 2018), abordamos as construções absolutas do português brasileiro<sup>29</sup>, seguindo a proposta de Mufwene (2008) a partir dessa perspectiva das margens. Nos termos dessa proposta, em situações de contatos de línguas (de dialetos, ou mesmo de idioletos) forma-se um banco de traços linguísticos, que compreende elementos lexicais, fonológicos, morfológicos, sintáticos, semânticos e pragmáticos das várias línguas em questão. Em suas interações comunicativas, os falantes selecionam alguns desses traços para construir seus enunciados na língua emergente, independentemente de qual seja a língua de que eles venham. Tendem a ser selecionados traços que são percebidos pelos falantes como congruentes entre si. Nossa hipótese para explicar as construções absolutas do português brasileiro foi a de que traços parciais de construções com deslocamento à esquerda do português do século XV foram escolhidos porque eles mostravam alguma congruência com alguns dos traços que caracterizam um tipo de construção do kimbundu, que tem sido chamada ‘passiva’ (Givón 2006). As construções absolutas do português brasileiro não são nem uma réplica de construções do português clássico, nem de construções do kimbundu. São antes construções inovadoras, que compartilham alguns traços com as línguas de superstrato e de substrato.

O exame das construções de inversão das línguas bantu que estamos fazendo neste texto vem agregar peso à contribuição que essa família de línguas africanas deu para o português brasileiro. Não só as construções absolutas parecem ter resultado da congruência parcial de traços entre construções de deslocamento à esquerda do português clássico e de uma das construções ‘passivas’ do kimbundu, como elas também se assemelham a sentenças de outras línguas bantu que aceitam, em posição pré-verbal, constituintes com papéis temáticos variados, que concordam com o verbo, como os exemplos entre (18) e (22) acima. As construções

<sup>28</sup> Para uma primeira abordagem do impacto de línguas indígenas da família tupi-guarani na constituição da gramática do português brasileiro, ver Negrão e Viotti (em preparação).

<sup>29</sup> Ver nota de rodapé 1 acima.

absolutas se diferenciam dessas na medida em que, além do participante do evento que é codificado na posição de sujeito, nenhum outro participante do evento é expresso na sentença, não chegando nem mesmo a ser semanticamente conceitualizado.

Essas construções de inversão das línguas bantu são também importantes para entender várias construções peculiares do português brasileiro, como aquelas apontadas por Pontes (1986), exemplificadas entre (10) e (14). Tanto quanto em várias línguas bantu, essas sentenças do português brasileiro têm um constituinte pré-verbal que, apesar de disparar a concordância com o verbo, tem um papel temático que não é aquele de agente (o protótipo da semântica do sujeito), podendo ser um locativo, um possuidor, um paciente, um beneficiário, um tema, etc. Para além disso, há um tipo de construção específica de inversão do sujeito em línguas bantu – a inversão de concordância default, cujos exemplos aparecem entre (23) e (26) – que apresenta grande similaridade com algumas das construções impessoais do português brasileiro, estudadas por Franchi *et al.* (1998), apresentadas entre (27) e (38). Há dois padrões que são seguidos tanto por muitas línguas bantu, quanto pelo português brasileiro: a sentença pode apresentar o verbo em primeira posição, sem concordância com o sujeito lógico da sentença que aparece em posição pós-verbal; ou, então, pode haver um sintagma locativo em posição anterior ao verbo, que tampouco concorda com o sujeito lógico que o segue.

A partir dessa mesma perspectiva marginal e inovadora sobre como o contato de línguas se dá, as línguas gbe parecem ter também contribuído para a emergência de mais uma peculiaridade do português brasileiro, já apontada desde Pontes (1986): a preferência pela codificação da estrutura informacional em detrimento da codificação estruturada das relações semânticas dos argumentos verbais. Pelo fato de o português brasileiro aceitar, em sua posição pré-verbal, constituintes sintáticos com semânticas variadas, Pontes enfatiza que, nessa língua, os falantes percebem o sujeito mais como tópico, do que como o participante do evento que tem papel temático de agente ou causa. Além disso, são muito comuns em português brasileiro construções como a exemplificada em (3), que são semelhantes àquelas em gungbe exemplificadas em (41): há em ambas um deslocamento de um constituinte para uma posição hierarquicamente mais alta que expressa o conteúdo informacional da sentença, com a presença de um pronome resumptivo na posição de onde esse constituinte saiu. Por fim, línguas gbe podem ser consideradas parte da explicação para a assimetria notada por Negrão (1997) relativa ao preenchimento ou não da posição de sujeito. Como visto, em algumas línguas gbe, como o ewe, há uma distribuição complementar entre pronomes logofóricos e pronomes não logofóricos, que é análogo ao comportamento de categorias vazias e pronomes plenos em português brasileiro.

Este nosso estudo sobre a emergência do português brasileiro como uma língua colonial sugere que devemos entender grande parte das novas línguas românicas como línguas híbridas (*cf.* Aboh 2015): elas naturalmente manifestam, em maior ou menor grau, características românicas, mas apresentam também semelhanças com muitas das línguas não indo-europeias faladas nos territórios conquistados e colonizados pelos europeus. Essas semelhanças, como já dissemos, são mais profundas do que o empréstimo de itens lexicais ou de estruturas morfossintáticas, e precisam ser investigadas a partir de uma análise cuidadosa das chamadas línguas de substrato. O estudo do contato de línguas, o de línguas emergentes que têm, em sua origem, a aquisição tardia de uma língua europeia por falantes nativos de outras línguas, os fenômenos que caracterizam línguas africanas e indígenas que as afastam do padrão indo-europeu, tudo isso tem ficado à margem da linguística nuclear. De modo geral, quando esses temas são trazidos para o núcleo, muitas vezes acabam por ser examinados a partir das perspectivas teórico-metodológicas concebidas para o estudo de línguas indo-europeias, criando

o risco de enviesar nosso entendimento sobre os fatos. Parece-nos que ouvir os linguistas que têm trabalhado na marginália da linguística pode ser uma estratégia promissora para a investigação de muitas particularidades das línguas mais recentes, no sentido de vir a fornecer-lhes um tratamento mais abrangente.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Aboh, Enoch O. 2004. *The morphosyntax of complement-head sequences. Clause structure and word order patterns in Kwa*. New York, Oxford University Press.
- Aboh, Enoch O. 2007. Leftward focus versus rightward focus: the Kwa-Bantu conspiracy., em *SOAS Working Papers in Linguistics*, 15: 81-104.
- Aboh, Enoch O. 2015. *The emergence of hybrid grammars*. Cambridge, Cambridge University Press.
- Aboh, Enoch O. 2020a. Lessons from neuro-(a)-typical brains: universal multilingualism, code-mixing, recombination, and executive functions, em *Frontiers in psychology*, 23 de abril de 2020, <https://doi.org/10.3389/fpsyg.2020.00488>
- Aboh, Enoch O. 2020b. Universal multilingualism: contact, acquisition and change, palestra apresentada no *Abralin ao Vivo-Linguists on Line*, 24 de junho de 2020.
- Altman, Cristina. 2011. A descrição da línguas ‘exóticas’ e a tarefa de escrever a história da linguística, em *Revista da ABRALIN*: 209-230.
- Ameka, Felix K. 1992. Interjections: the universal yet neglected part of speech, em *Journal of Pragmatics*, 18: 101-118.
- Ameka, Felix K. 2001. Ideophones and the nature of the adjective word class in Ewe, em F.K. Erhard Voeltz e Christa Kilian-Hatz (eds.), *Ideophones*. 19 de dezembro de 2001, <https://doi.org/10.1075/tsl.44.04ame>
- Ameka, Felix K. 2009. Information packing constructions in Kwa: micro-variation and typology, em Enoch O. Aboh e James Essegbey (eds), *Topics in Kwa syntax*. Dordrecht/ Heidelberg/ London/ New York, Springer: 141-176.
- Ameka, Felix K. 2017. Logophoricity, em A. Y. Aikhenvald e R.M.W. Dixon (eds.), *The Cambridge handbook of linguistic typology*. Cambridge, Cambridge University Press: 513-537.
- Ameka, Felix K. 2020. Linguistics marginalia and “grammars of use”, palestra apresentada no *Abralin ao Vivo-Linguists on Line*, 7 de julho de 2020.
- Ameka, Felix K. e David P. Wilkins. 2006. Interjections, em Jef Verschueren e Jan-Ola Östman (eds.), *Handbook of pragmatics*. 21 de dezembro de 2006, <https://doi.org/10.1075/hop.10.int12>
- Avelar, Juanito e Charlotte Galves. 2016. From European to Brazilian Portuguese: a parameter tree approach, em *Cadernos de Estudos Linguísticos*, 582: 237-256.
- Baker, Mark C. 2008. *The syntax of agreement and concord*. New York, Cambridge University Press.
- Bonvini, Emilio 2009. Revisiter, trois siècles après, Arte da lingua de Angola de Pedro Dias S.I. (1697), première grammaire du Kimbundu, em Margarida M. T. Petter e Ronald B. Mendes (eds.), *Exploring the African language connection in the Americas*. Proceedings of the Special World Congress of African Linguistics, São Paulo, Humanitas: 15-45.
- Chatelain, Henri Abraham. 1888-89. *Grammatica elementar do kimbundo ou língua de Angola*. Genebra, Typ. de Charles Schuchardt.
- Cobbinah, Alexander Yao. 2020. O contato linguístico transatlântico desde uma perspectiva africana, palestra apresentada na mesa-redonda ‘África e Brasil: Contatos Linguísticos, no *Abralin ao Vivo-Linguists on Line*, 11 de julho de 2020.
- Demuth, Katherine e Sheila Mmusi. 1997. Presentational focus and thematic structure in comparative Bant, em *Journal of African Languages and Linguistics*. 18: 1-19.
- Dingemans, Mark. 2017. On the margins of language: Ideophones, interjections and dependencies in linguistic theory, em N. J. Enfield (ed.), *Dependencies in language: on the causal ontology of linguistic systems*, Berlin, Language Science Press: 195-203.
- Duarte, Maria Eugenia Lamoglia. 2000. The loss of the ‘Avoid Pronoun Principle’ in Brazilian Portuguese, em Mary A. Kato e Esmeralda V. Negrão (eds), *Brazilian Portuguese and the null subject parameter*. Madrid, Vervuert-Iberoamericana: 17-36.
- Epps, Patience L. 2018. Contrasting linguistic ecologies: indigenous and colonially mediated language contact in northwest Amazonia, em *Language and Communication*, 62:156-169.

- Epps, Patience L. e Kristine Stenzel. 2013. Introduction: cultural and linguistic interaction in the Upper Rio Negro region, em Patience L Epps e Kristine Stenzel (eds.), *Upper Rio Negro. Cultural and linguistic interaction in Northwestern Amazonia*. Rio de Janeiro, Museu do Índio (FUNAI)/Museu Nacional: 13-50.
- Essegbey, James. 1999. Inherent complement verbs revisited: towards an understanding of argument structure in Ewe. *MPI series*, Wageningen, Ponsen e Looijen.
- Galves, Charlotte. 1993. O enfraquecimento da concordância no português do Brasil, em Ian Roberts e Mary Kato (orgs), *Português brasileiro: uma viagem diacrônica*. Campinas, Editora da UNICAMP: 387-408.
- Galves, Charlotte. 2001. A sintaxe do português brasileiro, em Charlotte Galves, *Ensaio sobre as gramáticas do português*. Campinas, Editora da UNICAMP: 43-59.
- Givón, T. 2006. Grammatical relations in passive clauses: a diachronic perspective, em W. Abraham e L. Leisiö (eds.) *Passivization and typology. Form and function*, Amsterdam/Philadelphia, John Benjamins: 337-350.
- Joseph, Brian D. 1997. On the linguistics of marginality: the centrality of the periphery, em *Chicago Linguistics Society*, 33: 197-213.
- Kato, Mary Aizawa 2015. Expletivos nulos e construções de tópico-sujeito no português brasileiro, em *Cadernos de Estudos Linguísticos*, 57: 7-21.
- Kato, Mary Aizawa e Maria Eugenia Lamoglia Duarte. 2018. Pre-verbal position in BP: a reinterpretation of the 'Avoid Pronoun Principle', em *Diadorim*, 20: 610 – 626.
- Kemmer, Susan. 1993. *The middle voice*. Amsterdam, John Benjamins.
- Krifka, Manfred. 1995. Swahili, em J. Jacobs; A. von Stechow; W. Sternefeld e T. Vennemann (eds), *Syntax. An international handbook of contemporary research*. Berlin/New York, de Gruyter: 1397-1418.
- Kula, Nancy C. e Lutz Marten. 2010. Argument structure and agency in Bemba passives, em Karsten Legère e Christina Thornell (eds.), *Bantu languages: analyses, description and theory* (East African Languages and Dialects, 20). Cologne, Rüdiger Köppe: 115-130.
- Lambrecht, Knud. 1994. *Information structure and sentence form: topic, focus, and the mental representation of discourse referents*. Cambridge/New York, Cambridge University Press.
- Lüpke, Friederike. 2016. Uncovering small-scale multilingualism, em *Critical Multilingualism Studies*, 4: 37-74.
- Lüpke, Friederike. 2020. The languages of multilingualism, palestra apresentada no *Abralin ao Vivo-Linguists on Line*, 11 de julho de 2020.
- Marten, Lutz. 2010. The great Siswati locative shift, em Anne Breitbarth, Christopher Lucas, Sheila Watts e David Willis (eds) *Continuity and change in grammar*. Amsterdam, Benjamins: 249-267.
- Marten, Lutz e Jenneke van der Wal. 2014. A typology of Bantu subject inversion, em *Linguistic Variation*, 14: 318-368.
- Modesto, Marcello. 2000a. Null subjects without 'rich' agreement, em Mary A. Kato e Esmeralda V. Negrão (eds), *Brazilian Portuguese and the null subject parameter*. Madrid, Vervuert – Iberoamericana: 147-174.
- Modesto, Marcello. 2000b. *On the identification of null arguments*. Tese de doutorado. University of Southern California. Los Angeles. [ em linha] Disponível em: [http://linguistica.fflch.usp.br/sites/linguistica.fflch.usp.br/files/inline-files/08\\_1.pdf](http://linguistica.fflch.usp.br/sites/linguistica.fflch.usp.br/files/inline-files/08_1.pdf)
- Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge/New York, Cambridge University Press.
- Mufwene, Salikoko S. 2003. Language endangerment: what have pride and prejudice got to do with it?, em Brian Joseph, Johanna Destefano, Neil G. Jacobs e Ilse Lehiste (eds.), *When languages collide. Perspectives on language conflict, language competition, and language coexistence*. Columbus, Ohio State University Press: 324-346.
- Mufwene, Salikoko S. 2008. *Language evolution. Contact, competition and change*. London/New York, Continuum International Publishing Group.
- Mufwene, Salikoko S. 2013. What African linguistics can contribute to evolutionary linguistics, em O. O. Orié e K. W. Sanders (eds.), *Selected Proceedings of the 43rd Annual Conference on African Linguistics*, Somerville, MA, Cascadia Proceedings Project: 52-67.
- Mufwene, Salikoko S. 2018. Population movements, language contact, linguistic diversity, etc. A postscript, em Ericka A. Albaugh e Kathryn M. de Luna (eds.), *Tracing language movement in Africa*. New York, Oxford University Press: 387-414.
- Negrão, Esmeralda V. 1992. Tem uma história que eu quero contar que começa assim: peculiaridades de uma construção existencial, em *Cadernos de Estudos Linguísticos*, 22: 81-90.
- Negrão, Esmeralda V. 1997. Asymmetries in the distribution of overt pronouns and empty categories in Brazilian Portuguese, em J. R. Black e V. Motapanyane, *Clitics, pronouns and movement*. Amsterdam/Philadelphia, John Benjamins: 217-255.
- Negrão, Esmeralda V. 1999 *O português brasileiro: uma língua voltada para o discurso*. Tese de Livre Docência. Universidade de São Paulo. São Paulo. Inédita.



- Negrão, Esmeralda V. e Evani Viotti. 2000. Brazilian Portuguese as a discourse-oriented language, em Mary A. Kato e Esmeralda V. Negrão (eds), *Brazilian Portuguese and the null subject parameter*. Madrid, Vervuert – Iberoamericana: 105-125.
- Negrão, Esmeralda V. e Evani Viotti. 2008. Estratégias de impessoalização no português brasileiro, em José Luiz Fiorin e Margarida M. T. Petter (eds.), *África no Brasil: a formação da língua portuguesa*. São Paulo, Contexto: 179–203.
- Negrão, Esmeralda V. e Evani Viotti. 2010. A estrutura sintática das sentenças absolutas no português brasileiro, em *Linguística*, 23: 37-58.
- Negrão, Esmeralda V. e Evani Viotti. 2011. Epistemological aspects of the study of the participation of African languages in Brazilian Portuguese, em Margarida M. T. Petter e Martina Vanhove (eds.), *Portugais et langues africaines. Études afro-brésiliennes*. Paris, Karthala: 13-44.
- Negrão, Esmeralda V. e Evani Viotti. 2014a. Contato entre quimbundo e português clássico: impactos na gramática de impessoalização do português brasileiro e angolano, em *Linguística*, 30: 289–330.
- Negrão, Esmeralda V. e Evani Viotti. 2014b. Brazilian Portuguese as a transatlantic language: agents of linguistic contact, em *Interdisciplinary Journal of Portuguese Diaspora Studies*, 3: 135–154.
- Negrão, Esmeralda V. e Evani Viotti. 2015. A investigação da estrutura argumental: por uma confluência dos instrumentais analíticos, em Luciana Storto, Bruna Franchetto e Suzi Lima (orgs), *Sintaxe e semântica do verbo em línguas indígenas no Brasil*. Campinas, Mercado de Letras: 261-285.
- Negrão, Esmeralda V. e Evani Viotti. (em preparação). Uniformitarianism and the emergence of the Brazilian variety of Portuguese.
- Pearson, Hazel. 2015. The interpretation of logophoric pronoun in Ewe, em *Natural Language Semantics*, 23:77–118
- Petter, Margarida M. T.; Esmeralda V. Negrão e Evani Viotti. 2018. The Africa-Brazil continuum: the case of passive and impersonal constructions, em Laura Álvarez Lopez, Perpétua Gonçalves e Juanito Avelar (eds.), *The Portuguese language continuum in Africa and Brazil*. Amsterdam/Philadelphia, John Benjamins: 211-235.
- Pilati, Eloísa e Rosana R. Naves. 2013. Desenvolvendo a hipótese da cisão da categoria pronominal no português brasileiro, em Denilda Moura e Marcelo Sibaldo (eds.), *Estudos e Pesquisas em Teoria da Gramática*. Maceió, Edufal: 233-253.
- Pilati, Eloísa, Rosana R. Naves e Heloísa M.M.L. Salles. 2017. On the syntax of subjects in Brazilian Portuguese: using the 'split' pronominal system as the basis for an alternative analysis, em *Diadorim*, 19: 99- 139.
- Quarezemin, Sandra e Anna Cardinaletti. 2017. Non-topicalized preverbal subjects in Brazilian Portuguese, compared to Italian, em *Annali di Ca' Foscari. Serie occidentale*, 51: 383-409.
- Pontes, Eunice S.L. 1986. *Sujeito: da sintaxe ao discurso*. São Paulo, Editora Ática.
- Rizzi, Luigi. 1997. The fine structure of the left periphery, em Liliane Haegeman (ed), *Elements of Grammar*. Dordrecht, Kluwer: 281-337.
- Rosa, Maria Carlota. 2013. *Uma língua africana no Brasil Colônia de Seiscentos. O quimbundo ou língua de Angola na Arte de Pedro Dias, S.J.* Rio de Janeiro, 7Letras.
- Santos da Silva, Wellington (em preparação) *A Língua Geral de Mina e a história do Ciclo do Ouro: estudo da formação de uma língua de contato*. Tese de Doutorado. São Paulo. Inédita.
- Sapir, Edward [1921] 1949. *Language. An introduction to the study of speech*. New York, Harcourt, Brace & World, Inc.
- Schleicher, August. 1989. *Darwinism tested by the science of language*. Londres, J.C. Hotten.
- Stenzel, Kristine e Khoo, Velda. 2016. Linguistic hybridity: A case study in the Kotiria community, em *Critical Multilingualism Studies*, 4: 75–110.
- Stringer, Chris. 2012. *Lone survivors. How we came to be the only survivors on earth*. New York, Times Books.
- Thomason, Sarah G. 2001. *Language contact: an introduction*. Washington, D.C, Georgetown University Press.
- Thomason, Sarah G e Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley, University of California Press.
- Viotti, Evani. 1999. *A sintaxe das sentenças existenciais do português do Brasil*. Tese de doutoramento. Universidade de São Paulo, São Paulo. Inédita.
- Viotti, Evani. 2017. Diário de campo. São Gabriel da Cachoeira, julho de 2016 e julho de 2017. São Paulo.
- Zwartjes, Otto. 2011. *Portuguese missionary grammars in Asia, Africa and Brazil, 1550-1800*. Amsterdam/Philadelphia, John Benjamins.