

A posição e a colocação de clíticos em predicados complexos: o português brasileiro visto a partir de duas vertentes

*Clitic position and clitic placement in complex predicates:
Brazilian Portuguese viewed from two trends*

Aroldo Leal de Andrade *

Universidade Estadual de Campinas, Campinas, São Paulo, Brasil

Zenaide de Oliveira Novais Carneiro **

Universidade Estadual de Feira de Santana, Feira de Santana, Bahia, Brasil

Resumo: Neste texto revisitamos a hipótese de acordo com a qual o português brasileiro é formado por duas vertentes – a norma culta e a norma vernacular ou popular –, a partir do problema da posição e da colocação de clíticos em predicados complexos. Para tanto, construímos uma base de dados composta por cartas escritas em diversos momentos dos séculos XIX e XX, classificadas quanto à escolaridade de seus remetentes. Dois grandes tipos de construções foram identificados: perífrases infinitivas e tempos compostos. Primeiramente, descrevemos a variação na subida de clíticos nos dados do século XIX, a fim de identificar os grupos de fatores relevantes, que se mostraram muito próximos aos dos encontrados no português europeu. Em seguida, analisamos a evolução quantitativa dos dados de posição e colocação até o século XX a partir da classificação em autores cultos e incultos, analisada a partir de um modelo de competição de gramáticas. Observamos que, quanto à posição dos clíticos, a mudança

* Pós-doutorando junto ao Departamento de Linguística do Instituto de Estudos da Linguagem da Universidade Estadual de Campinas – Unicamp, Campinas, São Paulo, Brasil; aroldo.andrade@gmail.com.

** Professora do Departamento de Letras e Artes da Universidade Estadual de Feira de Santana – UEFS, Feira de Santana, Bahia, Brasil; zenaide.novais@gmail.com.

foi paralela em ambas as vertentes estudadas. Isso sugere que a hipótese de que as vertentes formadas por falantes cultos e incultos tendem a se aproximar é uma generalização que depende da sensibilidade de cada fenômeno estudado a variáveis externas, sendo que fenômenos sintáticos como a posição de clíticos tendem a ser menos sensíveis a tais variáveis.

Palavras-chave: Português brasileiro. Predicados complexos. Colocação de clíticos. Mudança linguística. Variação diastrática.

Abstract: In this paper we revisit the hypothesis according to which Brazilian Portuguese is formed by two trends – the cultured norm and the vernacular or popular norm – with focus on the problem of the position and placement of pronominal clitics in complex predicates. To do so, we have constructed a database composed of letters written at different moments of the nineteenth and twentieth centuries, and classified according to the education level of their authors. Two structure types were identified: infinitival periphrases and compound tenses. First, we have described the variation in the application of clitic climbing in the nineteenth century data, in order to identify the relevant factor groups, which were very similar to those found for European Portuguese. Subsequently, we have analyzed the quantitative evolution of clitic position and placement until the twentieth century from the classification between cultured and non-cultured authors, with the help of a grammar competition model. We have noticed, regarding clitic position, that the change was parallel in the two studied trends. This suggests that the hypothesis that the trends formed by cultured and non-cultured speakers tend to come closer to each other is a generalization that depends on the sensitivity of each studied phenomenon to external variables, and that syntactic phenomena such as clitic position tend to be less sensitive to such variables.

Keywords: Brazilian Portuguese. Complex predicates. Clitic placement. Linguistic change. Diastratic variation.

1 INTRODUÇÃO¹

Este texto descreve a variação na posição e colocação de pronomes clíticos em predicados complexos do português brasileiro, considerando dados de falantes cultos e incultos.² Apesar de os clíticos dessa variedade do português já terem sido alvo de estudos baseados em diferentes tipos de *corpora* (cf. Pagotto, 1992; Galves, 2001; Lobo, 2001; Vieira, 2002; Schei, 2002, 2003; Carneiro, 2005; Martins, 2009, *inter alia*), o presente artigo pretende oferecer uma contribuição empírica e teórica para o estado da arte, pois compara dados de falantes do português brasileiro nos eixos diacrônico e diastrático, além de agregar informações desenvolvidas em teses recentes sobre a distribuição dos pronomes clíticos em predicados complexos em diferentes variedades do português.

Segundo Lucchesi (1994), a história do português brasileiro (doravante PB), pode ser sintetizada em termos de uma polarização entre normas vernáculas e normas cultas. Enquanto estas últimas seriam mais suscetíveis à pressão da gramática do português europeu moderno veiculada pela educação formal, as normas vernáculas expressariam sobretudo o PB vernacular, caracterizado por Mattos e Silva como resultante de uma “aquisição imperfeita em contexto de oralidade” (2004: 71). Dessa forma, especialmente ao longo do século XX, juntamente com os processos de industrialização, urbanização e vulgarização do ensino, seria observada uma aproximação entre essas vertentes, devido a sistemas de valores que subjazem aos padrões de uso.

O objetivo do trabalho é analisar tais resultados à luz da hipótese de Lucchesi (1994), segundo a qual haveria uma aproximação entre essas duas “vertentes” ou “grupos de normas”. Para tanto, partimos de uma base de dados composta por cartas escritas ao longo do século XIX e XX por remetentes com diferentes níveis de escolarização, informação obtida a partir de um cuidadoso trabalho filológico prévio sobre os autores das cartas. A abordagem adotada para a compreensão das distinções entre as vertentes baseou-se na noção de competição de gramáticas (cf. Carneiro, 2005; Carneiro, Galves, 2010[2006]; Martins, 2009).

1 Os autores reconhecem o apoio financeiro da Fapesp, via projeto n. 2011/19235-2 (Andrade), e do CNPq, via projeto n. 401433/2009-9, e da Fapesp, via projeto n. 5566/2010 (Carneiro). Este texto se relaciona ao processo Fapesp n. 2012/06078-9. Os autores agradecem a dois pareceristas anônimos e a Juliana S. Vianna pelos comentários a uma versão prévia deste trabalho. Desculpando-nos por eventuais erros ainda existentes, dedicamos este texto a Charlotte Galves, por seu incentivo em aliar a busca de dados de *corpora* com uma análise linguística teoricamente apurada, a fim de identificar diferentes gramáticas nos textos.

2 Consideramos o termo *inculto* como sinônimo de *não-culto*.

O texto está organizado da seguinte forma. Na seção 2 apresentamos alguns aspectos teóricos e metodológicos da pesquisa. Na seção 3 descrevemos a variação interna à gramática quanto à categoria que hospeda os clíticos, com base nos dados do século XIX em grupos verbais com as sequências de perífrases infinitivas (auxiliar + infinitivo) e de tempos compostos (auxiliar + gerúndio ou participípio), a partir do que concluímos que há evidências para se abordar a variação em termos de uma competição de gramáticas (cf. Kroch, 1989). Na seção 4 apresentamos os resultados relativos à categoria que hospeda o clítico (verbo auxiliar ou verbo lexical) e a colocação face a essa categoria, nos eixos diacrônico e diastrático, esta última considerando as duas vertentes (cultos e incultos). Na seção 5 discutimos as mudanças observadas, tendo em vista a hipótese de Lucchesi (1994). Finalmente, a seção 6 conclui o artigo.

2 ASPECTOS TEÓRICO-METODOLÓGICOS

Esta seção é dividida nas seguintes subseções: 2.1, sobre os pressupostos teóricos adotados; 2.2, sobre a composição da base de dados, e 2.3, sobre os procedimentos de análise estatística.

2.1 Pressupostos teóricos

Este trabalho parte de uma visão formal de concepção da linguagem, em que se diferencia língua-I(nterna) de língua-E(xterna). Em Chomsky (1986), tais conceitos se opõem da seguinte maneira: a língua-I se refere à faculdade da linguagem interna, individual de um ser humano, que codifica a intensionalidade, enquanto a língua-E consiste no produto cotidianamente exteriorizado em situações de uso, que manifesta a extensionalidade. A adoção de tais conceitos é necessária para que se observem padrões, não obstante a variação observada na língua-E.

Os trabalhos diacrônicos que adotam essa visão teórica têm de enfrentar a tensão entre essas duas concepções de linguagem, uma vez que os dados disponíveis são necessariamente aqueles da língua-E, a partir dos quais se busca chegar, indiretamente, à compreensão da língua-I. Segundo essa perspectiva, a variação interna ao falante pode ser enquadrada em pelo menos dois tipos:

- (i) *variação derivada da competição entre línguas-I.* Nesse caso, as opções seriam organizadas em gramáticas separadas presentes na mente/cérebro de um mesmo falante, devido ao fato de estarem presentes estímulos incompatíveis nos dados disponíveis à criança em fase de aquisição de língua, o que acaba gerando uma mudança morfossintática (cf. Kroch, 1989);

- (ii) *variação interna a uma mesma gramática*. Nesse caso, as opções em causa não são intrinsecamente incompatíveis, tendo em vista aspectos puramente formais, razão pela qual a variação é regida por requisitos de natureza funcional, ligados, por exemplo, ao processamento de sentenças ou à estrutura informacional.

Apesar da primazia conferida à abordagem formal, o sistema que acabamos de descrever é teoricamente coerente, na medida em que vários trabalhos têm reconhecido que a combinação de visões formais e funcionais é possível (e talvez preferível). Por exemplo, Chomsky (2005) reconhece a existência de um “terceiro fator” relevante para o *design* da linguagem humana, composto por princípios de domínio não-específico à gramática. Uma possível divisão de tarefas entre as duas visões é apresentada em Haspelmath (2000). Em suma, esses textos sugerem que não há propriamente uma oposição entre linguística formal e funcional, mas um *continuum* de teorias e práxis que pendem mais para uma ou outra visão.

Isso dito, nosso ponto de partida teórico nos leva a diferenciar entre dois fenômenos que, no geral, têm sido igualmente tratados na literatura sob o rótulo de “colocação”, “ligação”, ou mesmo “posicionamento” de pronomes clíticos: a categoria sintática que hospeda o clítico (posição), e a ordem do clítico face a essa categoria (colocação), como se observa nos exemplos seguintes (em que ‘cl’ indica *clítico*, ‘Aux’ indica auxiliar, ‘Lex’, o verbo *lexical*, e ‘=’, a relação de *cliticização morfossintática*):³

- (1) *Opções de posição de clíticos*
- a. clítico alçado: $cl=Aux\ Lex$ ou $Aux=cl\ Lex$
 - b. clítico não-alçado: $Aux\ cl=Lex$ ou $Aux\ Lex=cl$
- (2) *Opções de colocação de clíticos*
- a. próclise: $cl=Aux\ Lex$ ou $Aux\ cl=Lex$
 - b. ênclise: $Aux=cl\ Lex$ ou $Aux\ Lex=cl$

Tal distinção, estabelecida em Galves, Ribeiro, Torres Morais (2005) – exceto no que diz respeito ao uso explícito do termo posição –, deriva de uma repartição de tarefas entre sintaxe e morfologia: a *posição* é determinada no módulo sintático, enquanto a *colocação* é determinada no módulo morfológico (sendo a ênclise derivada a partir da próclise). Independentemente de assunções teóricas

3 A cliticização morfossintática inclui posição e colocação; a cliticização fonológica indica a relação com um hospedeiro que, não necessariamente, é o mesmo que se observa na sintaxe.

específicas, a empiria dá apoio à separação entre tais conceitos, pois, desde o trabalho seminal de Rizzi (1982[1976]), observou-se que é uma propriedade de certas línguas, como o italiano, o espanhol e o português europeu, que o pronome clítico possa ocorrer junto a um verbo auxiliar, e não junto ao verbo lexical não-finito que o seleciona, sendo esse processo opcional (cf. (3a-b)) e portanto de natureza distinta da questão do ordenamento do clítico face ao verbo, que no italiano é determinado pela flexão verbal (compare (3a) com (3c); e (3b) com (3d); exemplos ampliados a partir de Rizzi, 1982[1976], p. 1):

- (3) a. *Piero ti verrà a parlare di parapsicologia.* (posição: alçado; colocação: próclise)
 P. 2SG.DAT vir.3SG.FUT para falar.INF de parapsicologia
 b. *Piero verrà a parlarti di parapsicologia.* (posição: não-alçado; colocação: ênclise)
 c.**Piero verràti a parlare di parapsicologia.* (posição: alçado; colocação: ênclise)
 d.**Piero verrà a ti parlare di parapsicologia.* (posição: não-alçado; colocação: próclise)
 ‘Piero virá falar contigo sobre parapsicologia.’

Sempre que o clítico ocorre junto ao auxiliar, há uma evidência clara de que os dois verbos formam um predicado complexo, e o fenômeno responsável por essa posição (alta, em se considerando uma representação arbórea de baixo para cima) é chamado de *subida de clítico* (do inglês *clitic climbing*; ou opcionalmente, *alçamento de clítico*). Dada a inexistência de adjetivo cognato ao verbo *subir*, utilizamos os termos *clítico alçado* e *clítico não-alçado* para fazer referência às posições do clítico.

2.2 A composição da base de dados

A base de dados deste trabalho é composta por cerca de 550 sentenças com predicados complexos de “reestruturação”, uma construção que inclui certos verbos modais, aspectuais e temporais.⁴ Foram seis os *corpora* dos quais se extraíram os

4 O termo “reestruturação” é grafado entre aspas para indicar que é usado sem a adoção da análise original, ou seja, não se compartilha a ideia de que há um processo de reestruturação entre duas ou mais orações como requisito para a subida de clíticos. Na literatura, essa construção se distingue de outra, denominada “união de orações”, construção que envolve um grupo verbal composto por um verbo causativo ou perceptivo seguido de um verbo lexical no infinitivo (cf. Raposo, 1981, para a proposição do termo), não analisada neste trabalho devido ao pequeno número de dados disponíveis nos *corpora*.

dados, elencados a seguir, com suas respectivas abreviações e descrições, totalizando 1052 correspondências:⁵

- **Cartas Brasileiras (B)**: 500 cartas escritas entre 1809 e 1904, divididas em “cartas avulsas para vários destinatários” (1-208), “cartas para Severino Dantas” (209-310) e “cartas para Cícero Dantas Martins, o Barão de Jeremoabo” (311-500) (cf. Carneiro, 2005, volume 2);
- **Cartas do arquivo Dantas Jr. (D)**: 242 cartas dirigidas a João da Costa Pinto Dantas Jr., neto do Barão de Jeremoabo, escritas entre 1902 e 1962;
- **Cartas do arquivo Costa Pinto (P)**, 102 cartas trocadas por mulheres das famílias Araújo Pinho, Argolo, Carvalho, Costa Pinto, Ferreira de Moura e Wanderley, escritas entre 1911 e 1958;
- **Cartas em Sisal (S)**: 91 cartas recebidas por moradores das localidades de Riachão do Jacuípe, Conceição do Coité e Ichu, BA, entre 1906 e 2000, também denominadas “cartas de mãos cândidas do sertão baiano” (cf. Santiago, 2012);
- **Cartas da família Oliveira (O)**, 23 cartas escritas por dois membros da família Oliveira, radicada em Feira de Santana, BA, escritas entre 1962 e 1973;
- **Correspondências amigas do acervo de Valente, BA (V)**, 94 correspondências (cartas e cartões) escritas entre 1980 e 1993, dirigidas a Adelmário Carneiro Araújo.

Todos os dados de origem estão disponíveis para consulta e *download* dos interessados, em CD-ROM (Carneiro, 2011; Carneiro, Oliveira, Almeida, 2011; Carneiro, Santiago, Oliveira, 2011), e/ou num *corpus* eletrônico *online* (Corpus CE-DOHS, 2011, exclusivamente para os dois últimos *corpora*).

Os remetentes das cartas escritas no século XIX (*corpus B*) são diversos e foram divididos em dois grupos: os remetentes habitantes da costa, cultos; e os do interior baiano, semicultos. Já a escolaridade dos remetentes das cartas escritas no século XX foi classificada a partir do *corpus* de origem, tendo sido considerados cultos aqueles do *corpus D*, e incultos aqueles dos *corpora P, S, O* e *V*. Para essa classificação, considerou-se que falantes cultos são aqueles que alcançaram grau de instrução superior, como usual. Note-se que a composição dos *corpora* adotados contribui para a hipótese que se pretende verificar no presente trabalho, pois na classe dos incultos há progressiva escolarização durante o século XX: enquanto os remetentes do *corpus P* só tinham o diploma de primário (1.o ciclo do ensino fundamental, no atual sistema educacional brasileiro), os remetentes dos *corpora S* e *V*, que escreveram em anos posteriores, apresentam diferentes graus de

5 A base de dados está disponível aos interessados em: <https://dl.dropboxusercontent.com/u/1097038/Corpus_pred_complexos_2vertentes.xlsx>.

escolarização (ensino fundamental incompleto, ensino fundamental completo, ensino médio completo). Ou seja, espera-se que, quanto aos incultos, haja uma relativa aproximação face aos cultos em termos de escolarização. Uma visão geral da base de dados, cruzando a data de produção (excluindo correspondências não-datadas) com o tipo de amostra é mostrada na Figura 1.

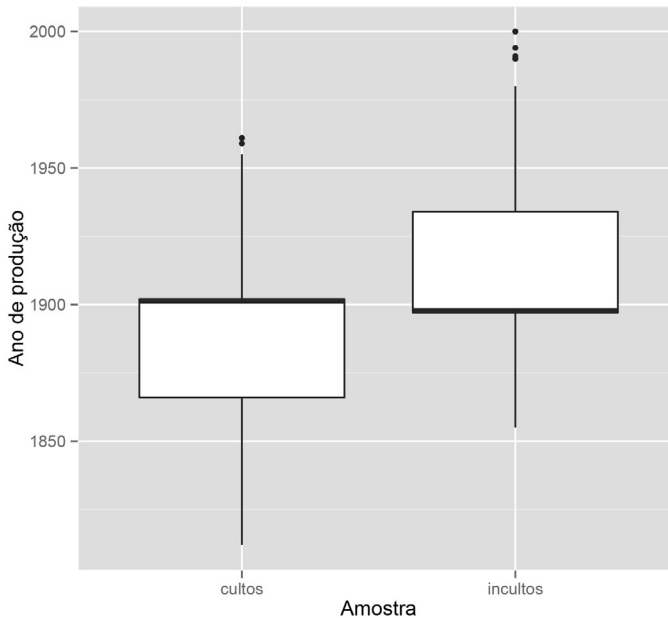


Figura 1. Distribuição dos dados por tipo de amostra e ano de produção

Essa figura demonstra que: os dados dos cultos se concentram no início do século XX, e se dispersam para a segunda metade do século XIX; os dados dos incultos se concentram no final do século XIX, e se dispersam para a primeira metade do século XX. Além disso, as amostras diferem quanto à sua extensão: os dados dos remetentes cultos praticamente não incluem a segunda metade do século XX; por outro lado, os dados dos remetentes incultos não incluem a primeira metade do século XIX. A obtenção de amostras mais simétricas entre si não é possível atualmente por causa da disponibilidade restrita de dados históricos do português brasileiro classificados quanto à escolaridade de seus produtores, o que é ainda mais difícil se se busca criar uma base de dados uniforme quanto ao gênero textual.

Ao lado da classificação por escolarização dos remetentes, os dados foram transversalmente separados quanto aos padrões de ocorrência da subida de clíticos no português europeu (PE): perífrases infinitivas e tempos compostos. Enquanto no contexto de perífrases infinitivas os clíticos podem ser alçados (4a-b) ou não (4a'-b'), no contexto de tempos compostos os clíticos são necessariamente alçados (5a-b), sendo o clítico não-alçado inaceitável (exemplos ampliados a partir de Brito; Duarte; Matos, 2003, p. 857):

(4) *Posições de clíticos em perífrases infinitivas (PE)*

- | | |
|--|---|
| a. O João não a vai provavelmente convidar. | a'não vai [...] convidá- la . |
| b. O João não a quer convidar. | b'não quer convidá- la . |

(5) *Posições de clíticos em tempos compostos (PE)*

- | | |
|--|---|
| a. O João tinha- a já convidado várias vezes. | a' . *... tinha já convidado- a ... |
| b. O João ia- se esquecendo do convite. | b' . *... ia esquecendo- se ... |

O agrupamento de (5a) e (5b) para o português brasileiro (PB) se baseia numa visão da linguística moderna, que prevê a existência de tempos perfeitos (*eu tinha/havia trabalhado muito*) e progressivos (*eu estou/andando trabalhando muito*) (cf. Perini, 2010, p. 238ss).⁶ Seguindo a divisão entre (4) e (5), a base de dados inclui 328 dados de perífrases infinitivas, e 235 dados de tempos compostos.

2.3 Métodos de análise dos dados

A variável dependente deste estudo consiste nas opções de uma matriz de possibilidades de posição e colocação de clíticos, acrescida da ambiguidade entre a ênclise ao verbo auxiliar ou a próclise ao verbo lexical, como ilustrado abaixo:

(6) *Fatores da variável dependente*⁷

6 Um dos pareceristas sugeriu que o agrupamento em (5) não se aplica aos dados da variedade brasileira, já que a variante em (5b') é admitida, inclusive por alguns gramáticos. Defendemos que o agrupamento é válido pois a construção é marginal, por ser geralmente limitada a clíticos de 3.ª pessoa, elementos que retêm a ênclise (cf. Galves, 2001, capítulo 8), por paralelismo com a colocação nas orações adjuntas gerundivas (p.ex., *Vendo-me doente, o João ficou.*). Mesmo que se questione isso, a classificação adotada permite uma separação ulterior entre grupos verbais participiais e gerundivos.

7 Os exemplos da base de dados são grafados seguindo a edição semidiplomática, excluindo-se a indicação de quebras de linha, porém mantendo-se a indicação de abertura de abreviaturas (em itálico) e de rasuras (em tachado). Para facilitar a leitura, indicamos em negrito os clíticos,

a. Próclise ao verbo auxiliar

Fico seriamente convencido que não **me** hade desamparar. [B 317]

b. Ênclise ao verbo auxiliar

Continua **se** a falar muito na retirada do *Ministro* da Fazenda... [B 261]

c. Próclise ao verbo lexical

... e comfio em sua generozidade que há de **mi**auxiliar no seguinte sentido. [B 445]

d. Ênclise ao verbo lexical

Hasde lembrar-**te** que te mandei um documento... [B 43]

e. Ambiguidade (ênclise ao verbo auxiliar ou próclise ao verbo lexical)

Poço **lhe** vender o grande branco por um conto de reis... [B 399]

Os dados classificados como ambíguos serviram para controle, tendo sido posteriormente descartados para fins da descrição da variação, razão pela qual tentamos desambiguá-los, tanto quanto possível. Como se trata de dados escritos, é importante considerar que a ambiguidade também pode advir das formas gráficas utilizadas nos manuscritos para indicar a cliticização. A base de dados inclui as seguintes possibilidades:

- para a próclise: o espaço (*me ha de faltar*) e a juntura na escrita (*ha de miauxiliar*).
- para a ênclise: o espaço (*Continua se a falar*) e a juntura com hífen (*Hasde lembrar-te*).

Uma vez que o espaço entre o clítico e forma verbal é possível nos dois tipos de colocação, essa é a opção que gera ambiguidade.⁸ Mesmo nesse caso, a desambiguação é possível em favor da próclise ao verbo lexical se houver um sintagma intercalado – cf. (7) –, ou uma preposição selecionada pelo auxiliar – cf. (8) –, ou se ocorrer um pronome forte em vez de um clítico, indicado na escrita pelo travamento silábico (*min* em vez de *me*, *lir* em vez de *lhe*) – cf. (9):

e em pontilhado os verbos do predicado complexo. Outros elementos que se queira realçar são identificados com sublinhado.

⁸ Agradecemos a um dos pareceristas por nos fazer esclarecer este ponto. Note-se que, nos corpora consultados, há outras formas indicativas de cliticização em predicados simples: o apóstrofe acompanhado de redução da vogal, no caso da próclise (*s'aclara*) e com o sinal de igualdade (*pedindo=lhe*) ou com o traço inferior (*pedir_lhe*), no caso da ênclise. Note-se que uma outra estratégia de desambiguação derivada da observação das formas verbais e pronominais não foi útil para nossa base de dados (considerando formas verbais terminadas em -s ou pronominais grafadas com -lo, -no e formas correlatas; cf. sugestão em Schei, 2003, p. 217).

- (7) Quero ainda uma vez lhe agradecer as boas e generosas palavras com que hontem eloquentemente saudou-me em nome do povo bahiano! [B 166]
- (8) Sinto mesmo, que não se apresente alguma vaga em concurso, por que elle não deixava de se inscrever, *muito* embóra, o governo o não nomeasse, más, ficaria todo o publico conhecendo de suas habilitações. [B 390]
- (9) conpadi sir u tenpos tirver bom min mandí Dizêr que eu quero lir mandar Dinheiro par u senhor min compar 20 casco di milho par mim [S 14]

A relativa produtividade de dados ambíguos na base de dados deve-se à opção (6c), a próclise ao verbo lexical, ser uma inovação do português brasileiro (cf. Pagotto, 1992, p. 152ss), irrestrita nos diversos tipos de predicados complexos, enquanto no português europeu ela é restrita a contextos em que há uma preposição selecionada pelo auxiliar, sendo necessariamente grafada com hífen nos exemplos correspondentes aos mostrados abaixo:

- (10) a. As esteiras agora foi que chegarão, em numero de 6, e com ordem de receberem 11\$500 do contrario expunhão avenda, achandoas caras só dei 10\$ e disse que hia lhe componçar, *oportador* aceitou, efição guardadas té sua ordem. [B 469]
- b. Como, porem, deixaste sem resposta as minhas primeiras cartas, a certeza de que, por esse meio, não tevi noticias directas tuas, tem me feito contacto com o que sei pelos amigos. [B 272]
- c. Por mim faço juizo dos outros, pois *somente* agora é *que vou me animando* com as providencias que tem tomado o Governo da União Federal. [B 438]

Além da variável dependente, e das variáveis utilizadas para situar diatriticamente e diacronicamente os dados da amostra (respectivamente, a data de produção da carta e a escolaridade do remetente), já discutidos previamente, foram codificados outros sete grupos de fatores como variáveis independentes, mostrados na Tabela 1.

Tabela 1. Variáveis independentes para a classificação dos dados

	Grupos de fatores	Fatores
a.	Verbo auxiliar	<i>(um fator para cada forma)</i>
b.	Verbo lexical	<i>gerúndio, infinitivo, participio</i>
c.	Sintagma interveniente	adverbial, nominal, preposicional; zero
d.	Preposição selecionada	<i>a, de, em; zero</i>
e.	Contexto sintático	<i>próclise obrigatória; ênclise obrigatória; variação 1 e variação 2</i>
f.	Forma do clítico	<i>(um fator para cada forma)</i>
g.	Função gramatical do clítico	<i>acusativo, dativo, dativo possessivo, dativo ético, inerente, reflexivo/recíproco, se anticausativo, se passivo, se indefinido, múltipla</i>

Utilizamos o software *Goldvarb X* primeiramente para a obtenção do padrão de variação quanto à posição do clítico, agrupando as opções em (6a-b) e (6c-d) e utilizando os dados do século XIX, que foram mais numerosos. Na análise binomial relacionada foram excluídos da contagem geral os dados de clíticos ambíguos quanto à variável dependente, ignorados os fatores categóricos face à variável dependente e agrupados os fatores semelhantes. Uma ilustração dos grupos de fatores possivelmente relevantes para a explicação da posição de clíticos é apresentada na Tabela 2.

Tabela 2: Agrupamentos de fatores para a análise binomial

	Grupos de fatores	Fatores
a'	Verbo auxiliar	- frequentes: <i>querer, poder, dever, ir, vir</i> ; - infrequentes: <i>haver (de), deixar (de), ter (de), começar (a), tornar (a), tornar (a), continuar (a), conseguir, costumar, esperar, desejar, pretender, ousar</i> . ⁹
e'	Contexto sintático	- próclise obrigatória; - ênclise obrigatória; - variação: <i>variação 1 + variação 2</i> . ¹⁰
f'	Forma do clítico	- 1. ^a pessoa: <i>me, nos</i> e correlatos (<i>mo, no-lo...</i>); - 2. ^a pessoa: <i>te, vos</i> e correlatos (<i>to, vo-lo...</i>); - 3. ^a pessoa anafórica: <i>o(s), a(s), lhe(s)</i> ; - 3. ^a pessoa não-anafórica: <i>se</i> .
g'	Função gramatical do clítico	- acusativos; - dativos argumentais; - dativos não-argumentais: <i>dativo ético, dativo possessivo</i> ; - clíticos de alteração de valência: <i>se passivo, se indefinido, se anticausativo</i> ; - clíticos exigidos pela forma verbal: <i>se reflexivo, se recíproco, se inerente</i> .

Segundo, para o estudo tanto da posição de clíticos quanto da colocação de clíticos – esse último agrupando-se as opções em (6a,c) e (6b,d) –, os resultados da variável dependente foram estudados face aos eixos diacrônico e diatrático, considerando ainda os subtipos de predicados complexos anteriormente referidos, sem levar em consideração todas as variáveis enunciadas acima.

9 A separação entre verbos frequentes e infrequentes baseou-se na divisão dos casos pelos tipos (*tokens e types*) na base de dados.

10 A distinção entre dois tipos de contextos de variação para a colocação de clíticos baseou-se na observação do português clássico (cf. Galves, Britto, Paixão de Sousa, 2005), em que há uma preferência pela próclise no contexto *variação 1* (verbo [auxiliar] precedido por elementos não-proclisadores, como sujeito, advérbio ou sintagma preposicional), e pela ênclise no contexto *variação 2* (verbo [auxiliar] inicial em segunda coordenada ou precedido por oração dependente). Vale lembrar que o português europeu moderno só apresenta os contextos de próclise obrigatória e de ênclise obrigatória, mais uma razão pela qual se deixaram os dados do século XX de lado, para a análise multivariada.

3 A DESCRIÇÃO DA VARIAÇÃO NA POSIÇÃO DE CLÍTICOS NO SÉCULO XIX

O objetivo desta seção é observar se a posição de clíticos nos dados do português brasileiro de épocas mais recuadas reflete a gramática do português europeu, que apresenta uma variação internamente motivada pela topicalidade do clítico e pelo grau de união oracional (cf. Andrade, 2010). Para tanto, utilizamos os dados produzidos no século XIX, por serem mais numerosos, não separando *a priori* os dados de acordo com a escolaridade dos remetentes, mas tão-somente de acordo com o subtipo de predicado complexo.

3.1 Perífrases infinitivas

Um resumo da análise binomial para os dados de perífrases infinitivas é apresentado na Tabela 3, em que se exibem os grupos de fatores em ordem decrescente de relevância para a variação na aplicação da regra de subida de clíticos, aferida a partir da amplitude dos pesos relativos. Do total de 243 dados disponíveis, foram analisados 221, com dados não-ambíguos.

Tabela 3. Análises multivariadas da contribuição de fatores selecionados como significantes para a probabilidade de subida de clíticos em perífrases infinitivas

<i>Português Brasileiro, século XIX</i>			
<i>Média corrigida = .448</i>			
<i>Resultado da estatística de comparação = -103.230</i>			
<i>N total</i>	221		
<i>Perífrases infinitivas</i>	<i>N</i>	<i>%</i>	<i>Peso relativo</i>
Função gramatical do clítico[†]			
Clíticos de alteração de valência	38 / 42	90.5	.87
Dativos não-argumentais	8 / 12	60.0	.58
Dativos argumentais	31 / 68	45.6	.51
Acusativos	18 / 54	33.3	.43
Clíticos exigidos pelo verbo	6 / 43	12.8	.13
<i>Amplitude</i>			<u>73</u>
Contexto sintático[‡]			
Contexto de próclise obrigatória	86 / 143	60.1	.63
Contexto de variação	11 / 65	16.9	.23
<i>Amplitude</i>			<u>40</u>
Frequência verbal			
Verbos frequentes	94 / 186	50.5	.55
Verbos infrequentes	7 / 35	20.0	.27
<i>Amplitude</i>			<u>28</u>
Escolaridade			
Cultos	69 / 126	54.8	.60
Incultos	32 / 95	33.7	.37
<i>Amplitude</i>			<u>23</u>

[†] Ignorados 2 dados de clíticos contraídos, por nocaute (subida categórica).

[‡] Ignorados 13 dados de clíticos em contexto de ênclise, por nocaute (não-subida categórica).

A relevância relativa dos grupos de fatores é semelhante à encontrada em Andrade (2010) para o português europeu. Em suma, nota-se que a subida de clíticos varia de acordo com dois grandes princípios funcionais: a topicalidade e a coesão estrutural. O primeiro princípio pode explicar os dois grupos de fatores mais relevantes, considerando a amplitude de seus pesos relativos:

- a topicalidade relacionada à função gramatical do clítico reflete o fato de que certas funções gramaticais codificam noções mais altas numa hierarquia de topicalidade (cf., entre outros, Silverstein, 1976);
- a topicalidade relacionada ao contexto sintático é mais alta se há um operador afetivo (cf. Klima, 1964), que corresponde a um elemento proclisador no português europeu, com escopo sobre o clítico.

O segundo princípio, a coesão estrutural, relaciona-se à frequência verbal porque verbos mais frequentes apresentam maior grau de gramaticalização, no sentido de auxiliar e verbo lexical se comportarem como uma unidade e assim facilitarem a subida de clíticos (cf. Davies, 1997, para um resultado semelhante para o espanhol). A variável nova aqui, e relativamente pouco relevante, consiste na escolaridade, em que os falantes incultos do PB se distanciam da gramática do português europeu, expressando clara preferência pela construção com o clítico não-alçado.

3.2 Tempos compostos

Uma análise binomial também foi aplicada aos tempos compostos, que apresentam variação quanto à posição de clíticos nos dados do PB, diferentemente do que ocorre no PE. Do total de 186 dados disponíveis no século XIX, somente 154 foram não-ambíguos entre subida e não-subida. O resultado dessa etapa é mostrado na Tabela 4.¹¹

¹¹ A possibilidade de ênclise aos gerúndios faz com que a percentagem de subida de clíticos com essa forma verbal seja um pouco menor que a encontrada com participios (77% contra 91%, respectivamente).

Tabela 4. Análises multivariadas da contribuição de fatores selecionados como significantes para a probabilidade de subida de clíticos com tempos compostos

<i>Português Brasileiro, século XIX</i>			
<i>Média corrigida = .90</i>			
<i>Resultado da estatística de comparação = -47.715</i>			
<i>N total</i>	<i>154</i>		
<i>Tempos compostos</i>	<i>N</i>	<i>%</i>	<i>Peso relativo</i>
Função gramatical do clítico[†]			
Clíticos de alteração de valência	21 / 22	95.5	.68
Dativos argumentais	65 / 70	92.9	.57
Acusativos	29 / 32	90.6	.50
Clíticos exigidos pelo verbo	15 / 23	65.2	.16
<i>Amplitude</i>			<i>52</i>
Contexto sintático			
Contexto de próclise obrigatória	108 / 121	89.3	[]
Contexto de variação	24 / 26	92.3	[]
Contexto de ênclise obrigatória	5 / 7	71.4	[]
<i>Amplitude</i>			<i>n/d</i>
Frequência verbal			
Verbos frequentes	114 / 125	91.2	[]
Verbos infrequentes	23 / 29	79.3	[]
<i>Amplitude</i>			<i>n/d</i>
Escolaridade			
Cultos	77 / 85	90.6	[]
Incultos	60 / 69	87.0	[]
<i>Amplitude</i>			<i>n/d</i>

[†] Ignorados 6 dados de dativos não-argumentais, e 1 dado de contraído, por nocaute (subida categórica).

Diferentemente do que se observou para as perífrases infinitivas, somente um grupo de fatores foi selecionado como significativo para a probabilidade de ocorrência da subida de clíticos em tempos compostos. Esse fator é a função gramatical do clítico, o mais relevante para as perífrases infinitivas. Esse resultado demonstra que os falantes parecem seguir um princípio de paralelismo, tratando de maneira semelhante os clíticos nessas duas construções, hipótese que se sustenta ao se observarem as porcentagens dos demais grupos de fatores na Tabela 4, que seguem as mesmas tendências mostradas na Tabela 3 (exceto uma ligeira alternância entre os contextos de variação e de próclise obrigatória, no grupo *contexto sintático*). O resultado da última tabela é crucial para demonstrar que há uma competição de gramáticas nos dois grupos (cultos e incultos), sendo a influência da gramática do PE mais relevante na mente/cérebro dos falantes cultos da amostra.

3.3 Discussão dos resultados

Em resumo, os resultados da variação na ocorrência da subida de clíticos apresentados nesta seção, com base em dados do português brasileiro do século XIX, indicam que nuances de variação internas à gramática do PE eram manejadas com certa maestria pelos falantes do português brasileiro estudados. Ou seja, pelo menos num estágio mais recuado da história do PB, uma explicação baseada somente na pressão normativa não é capaz de explicar a performance dos falantes (tanto cultos quanto incultos) com respeito às regras do PE, porque a gramática tradicional não havia descrito o fenômeno em toda a sua complexidade. Tal observação dá suporte à modelagem da variação interna ao falante seguida neste trabalho, e demonstra que estamos diante de um caso que combina os tipos (i) e (ii) mencionados na seção 2.1: é uma variação que se manifesta na competição entre línguas-I, e ao mesmo tempo no âmbito de uma mesma gramática.

Nesta altura temos uma motivação clara para estudar a variação na posição e colocação de clíticos, agora enfocando, no que tange à dimensão interna, a variação entre línguas-I, combinando-a com a dimensão externa, ou seja, a variação diastrática entre línguas-E (baseada no critério de escolaridade dos falantes, tendo por base informações sobre a sociohistória do português brasileiro). Cabe, no entanto, observar quais seriam as gramáticas (línguas-I) em competição. Os estudos sobre a questão em textos brasileiros são unânimes quanto à existência de três gramáticas em competição – do português brasileiro, português clássico e português europeu moderno –, observação inicialmente feita em Carneiro (2005) e elaborada em outros trabalhos.¹²

12 Conforme a autora: “A aparição de padrões convergentes de gramáticas distintas em compe-

Apesar dos trabalhos anteriores, considerando que estamos tratando somente dos predicados complexos, nossos resultados só podem se restringir à competição entre PB e PE, sendo este último termo aqui entendido como um conjunto de união do português clássico com o português europeu moderno. Primeiro, porque a posição de clíticos não apresenta mudança qualitativa entre o português clássico e o português europeu moderno, mas tão-somente qualitativa (decréscimo na percentagem de subida de clíticos; cf. Andrade, 2010, capítulo 4). Segundo, porque, mesmo para a colocação de clíticos, uma competição envolvendo três gramáticas torna-se mais difícil de ser comprovada, tendo em vista a pequena quantidade de dados de ênclise com clítico alçado do PE, sendo essa opção crucial para se postular a diferença entre português clássico e português europeu moderno.¹³ Esse último ponto se repete nos dados do PB, apresentados a seguir.

4 A COMPETIÇÃO DE GRAMÁTICAS EM DUAS VERTENTES

Nesta seção comparamos todos os dados quanto ao comportamento face à posição do clítico e à colocação (no domínio do verbo lexical), considerando os dois grandes tipos de construção considerados neste estudo, e a variável *escolaridade*. No que se refere à posição, enfocamos a variante conservadora, a subida de clíticos, e no que tange à colocação, enfocamos as opções inovadoras, a próclise ao verbo lexical e a posição/colocação ambígua.

4.1 A posição de clíticos

Mostramos os resultados da ocorrência da subida de clíticos para as perífrases infinitivas na Figura 2, nas duas vertentes (cultos e incultos), por metades de

tição nos textos escritos no Brasil, aliada aos resultados dessa pesquisa, com uma gramática tipicamente brasileira no século 19, mostraria não apenas a *competição de duas gramáticas, mas de três*” (Carneiro, 2005, pp. 16-17 (ênfase acrescentada)). No entanto, já em Pagotto (1992) há menção explícita da presença de diferentes gramáticas nos textos brasileiros: “após ter mudado em relação ao português clássico, [o português europeu moderno] passou a exercer pressão sobre o português brasileiro” (161). Uma ilustração crucial dessa afirmação encontra-se no uso da ênclise em orações afirmativas neutras, que muda, por exemplo, nos textos de duas constituições brasileiras, a de 1824 e a de 1892 (cf. Pagotto, 2013[1999]): a primeira refletiria sobretudo as gramáticas do português clássico e do PB, enquanto a segunda se aproximaria à gramática do português europeu moderno.

¹³ A interpretação da variação na sintaxe (colocação e posição) dos clíticos dos textos brasileiros como um caso de competição de gramáticas baseou-se crucialmente em trabalhos anteriores de sistematização dos dados de variação, iniciado em Lobo (1993) e também discutido e elaborado em Martins (1994), Torres Morais (1995), e Galves, Britto, Paixão de Sousa (2005), *inter alia*.

século. Nesta seção os dados não-ambíguos são considerados somente para a obtenção das percentagens, e a subida inclui as colocações proclítica e enclítica ao verbo auxiliar. Observa-se que os cultos apresentam um decréscimo paulatino do fenômeno, exceto por um pequeno acréscimo na primeira metade do século XX, com os seguintes percentuais: 62%-46%-51%-25%. Não muito diferente foi o resultado encontrado com os incultos, que também utilizaram cada vez menos a construção, sendo que uma queda mais pronunciada foi observada entre a segunda metade do século XIX e a primeira metade do século XX, com os percentuais: *n/d*-31%-5%-2%.

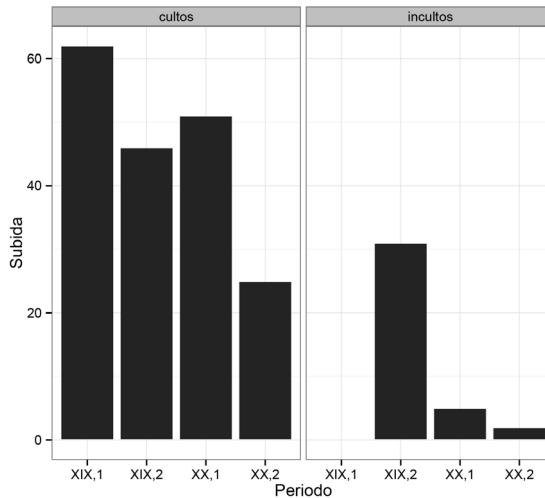


Figura 2: Ocorrência da subida de clíticos com perífrases infinitivas pelas duas vertentes do PB

Um resultado não muito diferente é observado para o contexto de tempos compostos, com a diferença de os percentuais de subida serem mais altos, como se vê na Figura 3. Para esse contexto, os remetentes cultos apresentaram decréscimo na ocorrência desse fenômeno entre a primeira e a segunda metades do século XIX, e não apresentaram dados na segunda metade do século XX (100%-82%-86%-*n/d*). Também houve decréscimo no grupo dos remetentes incultos (*n/d*-61%-35%-12%).

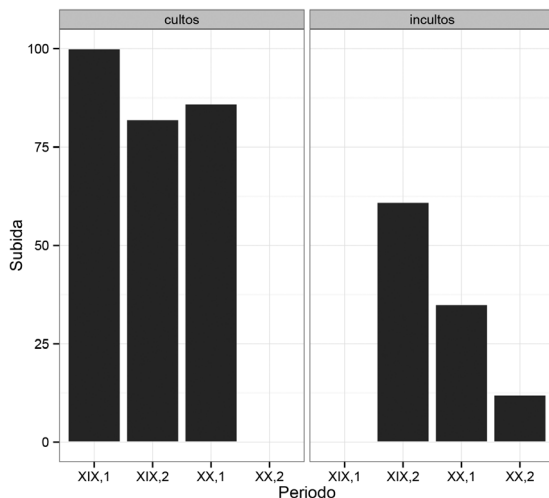


Figura 3: Ocorrência da subida de clíticos com tempos compostos pelas duas vertentes do PB

Em resumo, os resultados relativos à subida de clíticos demonstram que, para esse fenômeno, há uma aproximação entre as vertentes culta e inculta em decorrência de uma mudança numa mesma direção, que vem a ser a perda da subida de clíticos. O pequeno acréscimo de dados de subida com os cultos observado na primeira metade do século XX se deve, provavelmente, a uma mistura de dados de duas gerações, já que parte desses remetentes, que escreveram nos primeiros anos desse período, é da mesma geração dos que escreveram nos últimos anos do período anterior (fim do século XIX). Esse fenômeno demonstra que a competição de gramáticas é mais útil para a compreensão dessa mudança morfossintática do que tão-somente a realidade sociolinguística dos falantes, pois eles não seguem tendências opostas.

4.2 A colocação de clíticos

Nesta seção analisamos os dados de colocação de clíticos, primeiro observando as opções com os clíticos alçados, e em seguida com os clíticos não-alçados. Todas as porcentagens foram calculadas sobre o total de ocorrências dos fatores elencados em (6).

4.2.1 A colocação com clíticos alçados

Sempre que o clítico é alçado no português europeu, sua colocação segue as mesmas regras previstas para os predicados simples. Para os gráficos desta seção, agrupamos o contexto de variação na colocação de clíticos com o contexto de ênclise obrigatória, uma vez que essa foi a mudança ocorrida na transição entre o português clássico e o português europeu moderno, sem considerar o eixo diacrônico. Os gráficos na Figura 4 apresentam os resultados para as perífrases infinitivas e os tempos compostos.

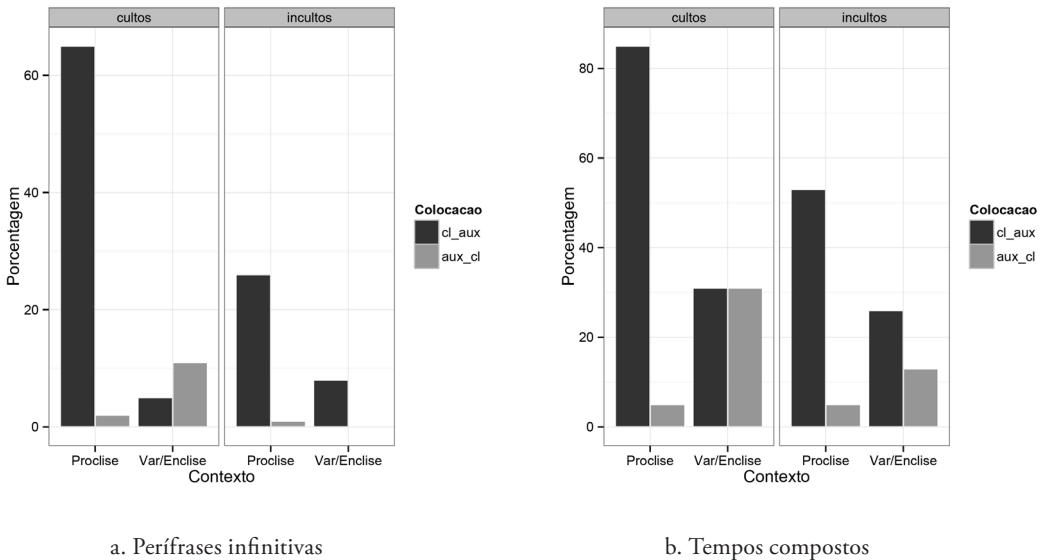


Figura 4. Opções de colocação de clíticos alçados quanto ao contexto sintático, considerando duas vertentes da sociohistória do PB

Os gráficos demonstram que as duas vertentes se comportam em paralelo em alguns aspectos, mas divergem em outros. No entanto, de forma geral, as semelhanças são maiores.

Começamos pelas perífrases infinitivas, com resultados mostrados na Figura 4a, em que os comportamentos são em geral paralelos (cultos – próclise: 65%-21%; var./ênclise: 5%-11%; incultos – próclise: 26%-1%; var./ênclise: 8%-0%). Por exemplo, ocorrem dados de ênclise em contexto de próclise obrigatória nos dois tipos de amostra, em pequeno número. Tais exemplos podem ser considerados hipercorreções, já que a ênclise é uma opção marcada no PB:

(11) *Perífrases infinitivas: ênclise com clítico alçado em contexto de próclise obrigatória*

a. ... que so poder-se-á verificar pelo balanço do fim do corrente. [B 289] (1902)¹⁴

b. ... e mesmo por isto quero te pedir que < ↑ PALAVRAS > procuramse distrair [O 16] (1968)

Da mesma forma, as duas vertentes se comportam de maneira bastante semelhante quanto aos dados nos demais contextos. Nem os remetentes cultos, nem os incultos apresentam dados de próclise em contexto de ênclise obrigatória (verbo em primeira posição). Além disso, ambos parecem seguir a gramática do português clássico até o começo do século XX, o que se observa pelo uso da próclise em contextos de variação:¹⁵

(12) *Perífrases infinitivas: próclise com clítico alçado em contexto de variação*

a. Seos Collegas da Marinha e da Justça lhe podem informar se tenho ou não serviços á Provincia... [B 68] (1861)

b. É verdade que o meo filho não é conhecido no magistério d'aqui, más lhe possô affirmar o seu preparo... [B 390] (1900)

A única diferença é que, nos poucos dados nesse contexto, os incultos preferem usar a próclise com clítico alçado, já nos dados do século XIX, nos contextos de variação, sendo esta substituída pela próclise com clítico não-alçado no começo do século XX, em ambas as vertentes. Interpretamos esse fato como indicativo de que os remetentes nos dois tipos de amostra sofreram influência da gramática do português clássico.

Passamos agora aos dados com tempos compostos, mostrados na Figura 4b. No geral, nota-se também um paralelismo quanto aos resultados quantitativos (cultos – próclise: 85%-5%; var./ênclise: 31%-31%; incultos – próclise: 53%-5%; var./ênclise: 26%-13%). Além disso, também se observam hipercorreções com esse grupo de dados, como a ênclise em contexto de próclise obrigatória, cf. (13):

14 Considerando-se a mesóclise como um subcaso da ênclise, como usual.

15 A única exceção, encontrada em período posterior pode ser considerada um caso isolado (o trecho em causa está dentro de uma oração intercalada:

(i) Ela será ali sua amiga dedicada e trabalhará, rigorosamente, [eu lhe possô assegurar], de acordo com a sua orientação. [D 170] (1949)

(13) *Tempos compostos (cultos e incultos): ênclise ao clítico alçado em contexto de próclise obrigatória*

a. O imperador saiu há dias na Nitheroy, com a qual está-se fazendo aqui tanto barulho quanto fizeram os portugueses com a rua Bertholomeu Zé Dias... [B 18] (1878)

b. Não melhor por que não estou-lhe yendo a todo momento. [S 57] (1977)

Por outro lado, somente os remetentes incultos generalizaram a próclise, inclusive em contexto de ênclise obrigatória (somente com o auxiliar *ser*, formador de voz passiva, em que a subida de clíticos é gerada pela gramática do PB):

(14) *Tempos compostos (incultos): próclise com clítico alçado em contexto de ênclise obrigatória*

a. **Mefoi intregui** a Intendencia com o debito de 90 \$ e o cofre vazio com 680 reis... [B 462] (1894)

b. **Mefoi intregue** pello Antero a sua presada carta de 1 do corrente acompanhada com 428.000 que fico recebido.[B 468] (1898)

Da mesma forma que se observou para os dados em perífrases infinitivas, após o começo do século XX não mais se observam casos de ênclise com clítico alçado (opção também pouco frequente nos dados do português europeu). Por conta disso reafirmamos que os dados não nos permitem observar uma influência do português europeu sobre a colocação de clíticos do PB (cf., no entanto, Carneiro, Galves, 2010[2006] e Martins, 2009, para trabalhos que exploram essa hipótese com base em dados de colocação em predicados simples).

4.2.2 A colocação com clíticos não-alçados

Após observar os dados de clíticos junto ao verbo auxiliar, podemos novamente mostrar os dados no eixo temporal, já que a colocação de clíticos nesse contexto não é sensível ao contexto sintático (e sim ao contexto fonético no PE, cf. Andrade, 2010, capítulo 2). Os gráficos a seguir apresentam as duas possibilidades de colocação, acrescidas da opção posição/colocação ambígua, considerando que na maior parte dos casos a ambiguidade corresponderá à próclise com clítico não-alçado.

Começamos pelas perífrases infinitivas, nas quais se observa que a correlação entre dados ambíguos e a próclise com clítico não-alçado (cl_lex), especialmente com os remetentes incultos, como se vê na Figura 5.¹⁶

16 As percentagens exatas mostradas na Figura 5 (na ordem ambíguo/cl_lex/lex_cl) são:

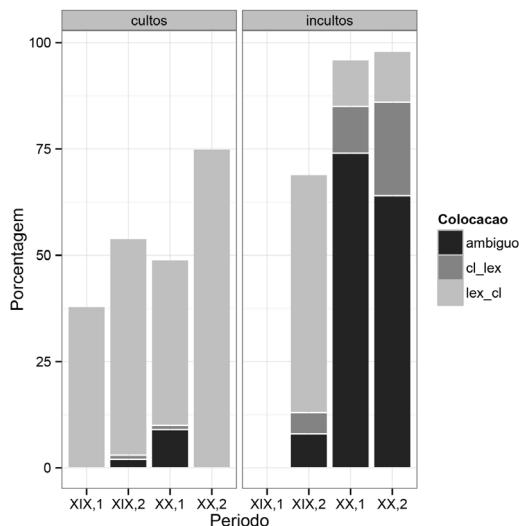


Figura 5. Opções de colocação de clíticos não-álçados em perífrases infinitivas, considerando duas vertentes da sociohistória do PB

Apesar de compartilharem a construção, as variantes apresentam uma diferença gritante, pois entre os cultos a ênclise predomina e, entre os casos de próclise, poucos são os que correspondente indubitavelmente à gramática inovadora – cf. (15) = (8) –, enquanto vários outros dados correspondem ao subcontexto que também é licenciado pela gramática do português europeu, em que há uma preposição *de* selecionada pelo verbo auxiliar – cf. (16):

(15) *Caso inovador (cultos): próclise com clítico não-álçado*

Quero *ainda uma vez* **lhe** agradecer as boas e generosas palavras com que hontem eloquentemente saudou-me em nome do povo bahiano! [B 166] (1882)

(16) *Caso conservador (cultos): próclise com clítico não-álçado*

O boato a meu respeito foi o preludio de alguma cousa que havia de **me** acontecer... [B 448] (1897)

- Cultos – XIX,1: 0%-0%-38%; XIX,2: 1%-1%-51%; XX,1: 74%-11%-11%; XX,2: 64%-22%-12%.

- Incultos – XIX,1: *n/d*; XIX,2: 8%-5%-56%; XX,1: 74%-11%-11%; XX,2: 64%-22%-12%.

Do lado dos incultos, a inovação se expressa em número bem maior, e igualmente sem elementos intervenientes, ou mesmo combinada com uma construção que hoje nos soa como tipicamente europeia (*ficar a + infinitivo*):

(17) *Caso inovador (incultos): próclise com clítico não-alçado*

a. ...pois o interessante é eu poder lhedizer com palavras tirada do coração que quanto mais dias passas mais aumenta o nosso amor [O 18] (1969)

b. Amanhã vai o Sr. Carvalho e fiça Maria a **nos** fazer boa companhia, sobretudo a Felipe. [P 6] (1922)

Passamos agora aos dados em tempos compostos, separados entre os participiais e os gerundivos, já que o PB, diferentemente do PE, apresenta diferenças quanto à colocação de clíticos não-alçados nesses contextos. Os resultados se veem na Figura 6.¹⁷

17 As percentagens exatas mostradas na Figura 6a (na ordem ambíguo/cl_lex/lex_cl) são:

- Cultos – XIX,1: *n/d*; XIX,2: 9%-3%-0%; XX,1: 5%-9%-2%; XX,2: *n/d*.
- Incultos – XIX,1: *n/d*; XIX,2: 24%-4%-4%; XX,1: 25%-17%-0%; XX,2: 60%-30%-0%.

Para a Figura 6b:

- Cultos – XIX,1: *n/d*; XIX,2: 14%-29%-0%; XX,1: 7%-0%-0%; XX,2: *n/d*.
- Incultos – XIX,1: *n/d*; XIX,2: 56%-0%-19%; XX,1: 64%-18%-9%; XX,2: 50%-33%-0%.

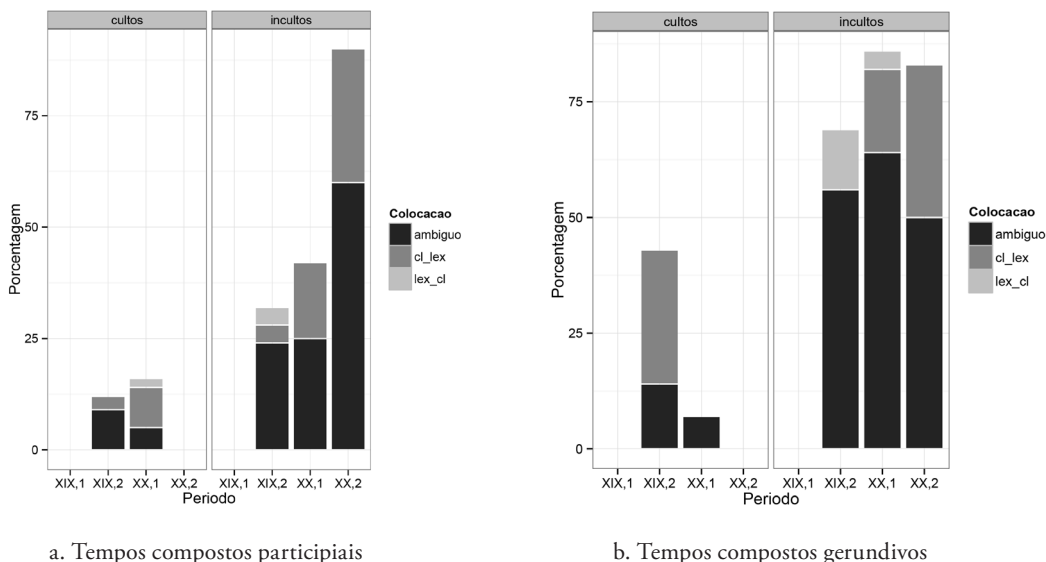


Figura 6. Opções de colocação de clíticos não-alçados em tempos compostos, considerando duas vertentes da sociohistória do PB

Devido à pequena quantidade de dados nessas construções, não se obtiveram resultados dos remetentes cultos para a primeira e última metades de século (XIX,1 e XX,2), nem dos remetentes incultos para a primeira metade do século XIX. Aqui a diferença principal diz respeito não à variação ênclise/próclise com os clíticos não-alçados, mas à variação de posição (entre subida e não-subida), sendo a não-subida preferida entre os remetentes incultos, por ser representante da gramática inovadora.

Além dessa observação geral, chamam à atenção os casos de ênclise com clíticos não-alçados (*lex_cl*), sempre em ocorrências próximas à transição entre os séculos XIX e XX. O fenômeno ocorre nas duas vertentes, no caso dos tempos compostos de participípio:

(18) Ênclise com clítico não-alçado: formas participiais

- a.** Ha poucos dias recebi carta do nosso amigo Vergne, que me diz estar pasan-do mal em Poços de Caldas, onde se acha, sendo provavel que já tenha transferido-se para Cambuqueira. [B 246] (1901)

b. Consta que no Rio tem havi[do] grande movimento dos E[studan]tes da Eschola militar e os Jacobi[nos] com a Policia e que alguns officiaes ou Generaes tem apre-
sentado-se em favor dos estudantes; [...] [B 368] (1895)

No caso dos tempos compostos de gerúndio, os dados se concentram entre os remetentes incultos:

(19) *Ênclise com clítico não-alçado: formas gerundivas*

a. Está demorando-se bastante, e toda demora é quase sempre prejudicial. [B 375] (1875)

b. Estou escrevendo-lhe esta debaixo de magna impressão e so no adiantado da noite é que procuro-o fazer. [B 453] (1897)

Apesar de essa última construção ser chancelada pela norma brasileira, ambas podem ser consideradas reflexos de uma hipercorreção, que surge a partir de um paralelismo efetuado pelos falantes entre os dados de tempos compostos e os dados de perífrases infinitivas. A sua ocorrência no começo do século XX parece decorrer da discussão denominada “questão da língua”, com efeitos sobre os falantes com certo grau de escolarização.

4.3 Discussão dos resultados

Os resultados desta seção demonstraram que há desenvolvimentos paralelos entre posição e colocação de clíticos, com especial realce para dados de hipercorreção, em que se expressam opções de posição/colocação vedadas na gramática do PE. Tal fato, por sua vez, pode nos auxiliar na compreensão do conceito de competição de gramáticas em tela.

No que se refere à posição dos clíticos, vimos que as duas vertentes apresentam queda nos índices de subida, apenas com maior rapidez entre os incultos. A única diferença diz respeito aos tempos compostos, em que falantes cultos por vezes ainda usam a subida (*Ele me tinha visto*). No entanto, o uso característico do PB é a próclise com clítico não-alçado, exceto com o tempo composto formado por *ser* + particípio, usado para formar a voz passiva (cf. Pagotto, 1992).¹⁸ Ainda assim, nota-se que, como o argumento interno acusativo

18 Outros dois contextos são:

- clíticos éticos (*Não me venha fazer bagunça*) e indefinidos (*Não se pode entrar na igreja de bermuda*) em contexto de “reestruturação” com perífrases infinitivas;

do verbo é expresso como sujeito, o clítico deve corresponder a um argumento interno dativo (*Isso me foi dado*).

Quanto à colocação de clíticos, vimos que não há como diferenciar uma competição entre três gramáticas (PB, português clássico e português europeu moderno) no caso dos dados em tela, já que não há dados de ênclise com clíticos alçados depois do começo do século XX e, até então, os dados refletem mais claramente a gramática do português clássico, em que havia um contexto de variação. Com os clíticos não-alçados, no caso das perífrases infinitivas os remetentes cultos apresentam uso cada vez maior da ênclise ao verbo lexical; já no caso dos tempos compostos, eles mantêm algum uso de clíticos alçados, seguindo em ambos os casos a gramática do português europeu. Isso contrasta com os resultados dos remetentes incultos, que progressivamente usam a opção inovadora.

Os fenômenos de hipercorreção estudados demonstram a relevância da pressão normativa no *design* dos dados. A única forma de levar esse fator em consideração passa por considerar que a competição de gramáticas em tela deriva da aquisição de uma gramática idealizada via escolarização, concepção que contrasta com aquela proposta originalmente em Kroch (1989), em que o que está em jogo são duas gramáticas advindas de línguas-I numa situação de contato (cf. os termos “competição de gramáticas *na escrita*” em Martins, 2009, e “competição de gramáticas *nos textos*” em Carneiro, Galves, 2010[2006] (ênfase acrescentada)).

Tendo em vista os resultados de variação vistos na seção anterior, vimos que uma influência da gramática do PE é esperada pelo menos no caso dos falantes que tiveram dados diretos de PE à sua disposição, durante o processo de aquisição de linguagem. Ou seja, temos aqui o efeito de uma competição entre línguas-I, que é gradativamente substituído por uma competição entre língua-I e “periferia” da gramática, para utilizar um termo de Chomsky (1981). Esse quadro nos faz esperar que o fator diastrático, externo, tenha importância cada vez maior na manutenção de distinções entre as duas vertentes, à medida que passa o tempo, em vez de fatores internos ao falante. O que surpreende, no entanto, é o grande paralelismo observado entre as duas vertentes (à exceção da

- clíticos correspondentes ao argumento externo do infinitivo em construções de marcação de Caso excepcional com verbos causativos e perceptivos (*Ele me mandou/viu entrar*).

Esses casos não envolvem movimento do clíticos para o domínio infinitivo, sendo, portanto, casos espúrios de subida derivados por características específicas dos clíticos em tela (cf. Andrade, 2010: capítulo 5, para algumas observações relacionadas). Já a permanência da subida de clíticos com *ser* + participípio pode ser explicada pela dinâmica da própria voz passiva (cf. Reis, 2011 e Andrade, 2013, que discorrem sobre o problema no âmbito de análises diferentes para a perda da subida de clíticos no PB; sobre essa última questão, cf. também Cyrino, 2010).

manutenção artificial de alguns padrões, entre os remetentes cultos). Passamos a analisar essa questão em seguida.

5 A EVOLUÇÃO DIACRÔNICA DOS DADOS E A HIPÓTESE DE LUCCHESI (1994)

A questão principal que nos propomos a responder diz respeito à hipótese levantada em Lucchesi (1994), de que haveria uma tendência de aproximação entre as normas culta e vernacular, identificadas como duas vertentes de um diassistema que, para o autor, consistiria no próprio português brasileiro, numa perspectiva sociolinguística. Nesse texto a hipótese é proposta ao mesmo tempo como uma agenda de pesquisas, pois o autor se baseou inicialmente em considerações sobre a sociohistória do PB, mais do que sobre dados quantitativos próprios. Nesse sentido, buscamos contribuir para verificar a extensão em que essa suposição é verdadeira.

Ao final do texto de Lucchesi (1994) a hipótese de aproximação entre as duas vertentes do PB é logo explicada em termos de outra hipótese, a de *crioulização prévia* do PB vernacular (cf. Guy, 1989), posteriormente reformulada em Lucchesi (2003) em torno da noção de *transmissão linguística irregular* do PB, com a consequência prática de não se postular um crioulo na base de (todos) os falares populares dessa variedade do português. Com base nesse arcabouço teórico, Lucchesi (1994) critica a posição que intui de trabalhos como Tarallo (1993), em que o PB é referido como “um todo, constituindo uma realidade variável, mas unitária” (p. 25). Finalmente, Lucchesi (1994) conclui que os resultados de Tarallo sobre fenômenos morfossintáticos como a realização do sujeito e do objeto, os quais apontariam para um afastamento do PB face ao PE, seriam devidos à consideração de dados de padrões linguísticos característicos dos segmentos escolarizados, dos falantes das classes média e alta, ou seja, o que se passa no âmbito do que ele, Lucchesi, denomina norma culta. Lucchesi chega a atribuir isso a uma maior inclinação de F. Tarallo para a teoria da gramática, deixando problemas cruciais da abordagem sociolinguística sem um tratamento satisfatório.

Alguns resultados quantitativos claros para a hipótese de aproximação das duas vertentes, culta e vernacular, são apresentados em outro trabalho, Lucchesi (2001), em que novamente se faz referência a dados relativos a fenômenos de concordância. O autor apresenta o aumento dos níveis de concordância a partir de estudos em tempo aparente, segundo os quais tanto a concordância verbal quanto a nominal são mais frequentes nos falantes de menor faixa etária (cf., entre outros, os dados coletados pelo autor e colaboradores, à p. 119). Nessa altura, Lucchesi volta-se à sua outra linha argumentativa, sobre o fenômeno da transmissão irregular do PB, não tocando mais no problema da aproximação das vertentes.

Em face de tudo o que foi discutido nesses trabalhos, cabe questionar como outros fenômenos morfossintáticos, como a posição e colocação de clíticos, se comportam. O que se viu na seção anterior é que a separação dos dados em duas vertentes constituintes da sociohistória do PB não dá suporte à hipótese apresentada em Lucchesi (1994), e sim a uma abordagem mais próxima à que se defende em Tarallo (1993). Em suma, ambas as vertentes se distanciam da gramática do PE, sendo distintas somente quanto à rapidez com que ocorre a mudança.

No entanto, não acreditamos que esse resultado invalide a intuição exposta em Lucchesi (1994), mas que demonstra que a extensão de suas previsões é mais limitada do que ele previa originalmente. A origem do problema é que F. Tarallo e D. Lucchesi analisavam, em larga medida, fenômenos de natureza distinta, com o seguinte comportamento diacrônico: *fenômenos relacionados à sintaxe* (estratégias de relativização, preenchimento de sujeito/objeto) *parecem mudar na mesma direção*, enquanto *fenômenos relacionados à morfologia* (concordância verbal, concordância nominal) *parecem mudar em direções opostas, com tendência à aproximação*. Não outra pode ser a nossa conclusão, senão a de que fenômenos sintáticos estão menos sujeitos à variação devida a fatores de ordem externa, se comparados a fenômenos morfológicos (e fonológicos). De todos esses fenômenos, o domínio correto de traços flexionais (de caso, concordância, tempo, entre outros) é uma das tarefas mais difíceis na aquisição de língua.

Voltando à questão da distinção entre posição e colocação de clíticos, havíamos dito que o primeiro fenômeno seria regido pela sintaxe, e o segundo, pela morfologia, seguindo Galves, Ribeiro, Torres Morais (2005). Isso nos leva a questionar: não seria o caso de haver uma diferença entre posição e colocação, no sentido das hipóteses quantitativas que acabamos de mencionar? Em outras palavras, não é de se esperar que a colocação de clíticos nas duas vertentes apresentem uma tendência à aproximação? Os dados presentes nessa pesquisa não nos permitem desenvolver a questão, já que os contextos sintáticos para a colocação de clíticos determinam tão-somente a posição alta, e mesmo no PE a ênclise com clítico alçado é muito rara. No entanto, dados de outros autores que consideraram a colocação em predicados simples podem nos auxiliar nessa tarefa, considerando o contexto de ênclise obrigatória:

- Em Schei (2002, p. 63; 2003, p. 159), dois estudos baseados em dados de romances brasileiros, se observa que a próclise inovadora é quase inexistente com o verbo em posição inicial (V1) no século XIX (1%, ou seja, 10 casos num total de 961 dados), enquanto que se eleva a um quarto dos dados no século XX (25,8%, ou seja, 120 casos em 465 dados);

- Em Martins (2009, p. 168), estudo baseado em peças de teatro catarinense dos séculos XIX e XX, a próclise com V1 corresponde a 3% (16/566) dos dados no século XIX, e 32% (91/265) no século XX.

Os estudos acima se referem à colocação de clíticos com falantes cultos, sendo a percentagem de próclise inovadora maior nas peças de teatro porque nelas tende-se a imitar ou incluir exemplos de fala não-monitorada. Do lado dos falantes incultos encontramos o seguinte:

- Em Vieira (2002, p. 202ss) se demonstra, com base em dados orais, que os falantes jovens usam mais a próclise do que os falantes idosos, inclusive no contexto de ênclise obrigatória (94% contra 86%).

Ou seja, posição e colocação parecem mudar na mesma direção em ambas as vertentes. No entanto, o sentido conferido aos “fenômenos relacionados à morfologia” mencionado acima pode ser correto se se consideram entre os últimos somente aqueles que estão ligados à realização de morfemas gramaticais (gramemas) na língua. Nesse sentido, o estudo sobre os predicados complexos aqui realizado revela um contexto de maior dificuldade para o aprendiz de língua, o que fica demonstrado pela ocorrência de hipercorreções que, no entanto, não se limitam aos falantes incultos, se bem que ocorram com maior frequência entre esses.

6 CONSIDERAÇÕES FINAIS

Neste texto analisamos os padrões de posição e colocação de clíticos em predicados complexos do português brasileiro. A partir de estudo baseado em *corpora* de correspondências pessoais, demonstramos que a separação dos dados é útil para identificar os diferentes padrões de mudança em duas vertentes formadoras do diassistema do PB. O estudo partiu de uma modelagem em termos da proposta de competição de gramáticas, considerando o manejo, pelos falantes, de fatores internos muito refinados, relativos à variação entre subida e não-subida de clíticos. Após a análise dos dados, concluímos que a previsão feita em Lucchesi (1994) não se confirma para o caso da posição de clíticos, uma vez que não se verifica um compartilhamento de traços entre PB culto e PB vernacular, e sim uma mudança generalizante para as duas vertentes, com maior velocidade no caso dos falantes incultos. No caso da colocação de clíticos também intuímos, a partir de dados de trabalhos que analisaram o fenômeno em predicados simples, que a mudança tampouco parece ter ocorrido segundo a previsão do autor. A diferença face aos fenômenos de concordância diz respeito à sua maior dependência à morfologia, e

mais especialmente à realização de morfemas gramaticais, que seria mais sensível à variação devida a fatores de ordem externa.

O trabalho também demonstrou que, para o fenômeno estudado, a gramática do português brasileiro advém, pelo menos em parte, da gramática do português clássico, que pode ser caracterizado como um ponto de bifurcação entre PB e PE (cf. Paixão de Sousa, Kewitz, 2011). A competição do PB com a gramática do português europeu moderno se dá mais propriamente no sentido de uma competição entre língua-I e a “periferia” da gramática, que contém dados advindos da escolarização. Finalmente, os resultados servem como um convite à construção, análise e disponibilização de *corpora*, especialmente do português brasileiro popular, com informações precisas sobre a origem de seus falantes, de tal forma a verificar o critério de separação entre sintaxe e morfologia aqui proposto, no que tange à hipótese lançada em Lucchesi (1994).

REFERÊNCIAS

Andrade AL. A subida de clíticos em português: um estudo sobre a variedade europeia dos séculos XVI a XX [tese]. Campinas: Universidade Estadual de Campinas; 2010.

Andrade AL. Uma abordagem unificada para a posição de clíticos em português brasileiro e em francês. *Estudos da Língua(gem)*. 2013;11(2):63-81.

Brito AM, Duarte I, Matos G. Tipologia e distribuição das expressões nominais. In: Mateus MHM, Brito AM, Duarte I, Faria I. *Gramática da Língua Portuguesa*. 5. ed. rev. aum. Lisboa: Caminho; 2003. p. 795-867.

Carneiro ZON. *Cartas Brasileiras (1809-1904): um estudo linguístico-filológico* [tese]. Campinas: Universidade Estadual de Campinas; 2005.

Carneiro Z. *Cartas Brasileiras (1809-2000): coletânea de fontes para o estudo do português*. vol. 1 [CD-ROM]. Feira de Santana: UEFS; 2011.

Carneiro Z, Galves C. Variação e Gramática: colocação de clíticos na história do português brasileiro. *Revista de Estudos da Linguagem*. 2010; 8(2):7-38. [Comunicação original: Carneiro Z, Galves C. Clitic placement in the history of Brazilian Portuguese: a case of three-grammar competition. Ms. Diachronic Generative Syntax Conference 9, Trieste; 2006. Acesso: 10 jan. 2014. Disponível em: <<http://www2.units.it/digs9/pdf/carneiro&galves.pdf>>.

Carneiro ZON, Oliveira MF, Almeida NLF. Cartas brasileiras (1809-2000): coletânea de fontes para o estudo do português. vol. 2 [CD-ROM]. Feira de Santana: UEFS; 2011.

Carneiro ZON, Santiago HS, Oliveira K. Cartas brasileiras (1809-2000): coletânea de fontes para o estudo do português. vol. 3 [CD-ROM]. Feira de Santana: UEFS; 2011.

Chomsky N. Lectures on government and binding: the Pisa Lectures. Dordrecht: Foris; 1981.

Chomsky N. Knowledge of language: its nature, origin and use. New York: Praeger; 1986.

Chomsky N. Three factors in language design. *Linguistic Inquiry*. 2005;36(1):1-22.

Carneiro ZON, Oliveira MF. Corpus CE-DOHS. Corpus Eletrônico de Documentos Históricos do Sertão (Fapesb 5566/ 5566/2010/Consepe/UEFS: 202/2010) [banco de dados na Internet]; 2011. Feira de Santana: Universidade Estadual de Feira de Santana. Acesso: 10 fev. 2013. Disponível em: <www.uefs.br/cedohs>. Acesso: 10 fev. 2013.

Cyrino S. On romance syntactic complex predicates: why Brazilian Portuguese is different. *Estudos da Língua(gem)*. 2010;8(1):187-222.

Davies, M. A corpus-based approach to diachronic clitic climbing in Portuguese. *Hispanic Journal*. 1997;17:93-111.

Galves C. Ensaios sobre as gramáticas do português. Campinas: Editora da Unicamp; 2001.

Galves C, Britto H, Paixão de Sousa MC. The Change in clitic placement from classical to modern European Portuguese: results from the Tycho Brahe Corpus. *Journal of Portuguese Linguistics*. 2005;4(1):39-67.

Galves C, Ribeiro I, Torres Morais MA. Syntax and morphology in the placement of clitics in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*. 2005 4(2):143-177.

Guy G. On the nature and origins of Popular Brazilian Portuguese. In: *Estudios sobre el Español de América y Lingüística Afroamericana: ponencias presentadas en el 45.º*

Congreso Internacional de Americanistas. Bogotá: Instituto Caro y Cuervo; 1989. p. 226-244.

Haspelmath M. Why can't we talk to each other? A review article of 'Newmeyer, Frederick. 1998. Language form and language function. Cambridge: MIT Press.'. *Lingua* 2000;100(4):235-255.

Klima ES. Negation in English. In: Fodor JA, Katz JJ, editors. *The Structure of Language: Readings in the Philosophy of Language*. Englewood Cliffs: Prentice Hall; 1964. p. 246-323.

Kroch A. Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change*. 1989 1(3):199-244.

Lobo T. A colocação dos clíticos em português: duas sincronias em confronto [dissertação]. Dissertação. Lisboa: Universidade de Lisboa; 1993.

Lobo T. Para uma sociolinguística histórica do português do Brasil: edição filológica e análise linguística de cartas particulares do recôncavo da Bahia, século XIX [tese]. São Paulo: Universidade de São Paulo; 2001.

Lucchesi D. Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa*. 1994;12:17-28.

Lucchesi D. As duas grandes vertentes da história sociolinguística do Brasil (1500-2000). *D.E.L.T.A.* 2001;17(1):97-130.

Lucchesi D. O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In: Roncarati C, Abraçado J, organizadores. *Português Brasileiro: contato linguístico, heterogeneidade e história*. Rio de Janeiro: 7 Letras; 2003. p. 272-284.

Martins AM. Clíticos na história do português [tese]. Lisboa: Universidade de Lisboa; 1992.

Martins MA. Competição de gramáticas do português na escrita catarinense dos séculos 19 e 20 [tese]. Florianópolis: Universidade Federal de Santa Catarina; 2009.

Mattos e Silva RV. De fontes históricas para a história social linguística do Brasil: em busca de indícios. In: *Ensaios para uma sócio-história do português brasileiro*. São Paulo: Parábola; 2004. p. 69-90.

Pagotto E. A posição dos clíticos em português: um estudo diacrônico [dissertação]. Dissertação. Campinas: Unicamp; 1992.

Pagotto EG. A norma das constituições e a constituição da norma no século XIX. *Letra*. 2013;8(1-2):31-50. [Comunicação original: Pagotto, EG. A norma das constituições e a constituição da norma no século XIX. Ms. III Seminário do Projeto Para a História do Português Brasileiro. Campinas; 1999].

Paixão de Sousa MC, Kewitz V. Vésperas Brasileanas: uma agenda para os estudos sintáticos do Português Brasileiro nos primeiros séculos. *Revista Portuguesa de Humanidades*. 2011;1:67-92.

Perini MA. *Gramática do Português Brasileiro*. São Paulo: Parábola; 2010.

Raposo ES. A Construção ‘União de Orações’ na Gramática do Português [tese]. Lisboa: Universidade de Lisboa; 1981.

Reis FEB. A perda da subida de clíticos no Português Brasileiro: séculos XIX e XX [dissertação]. Campinas: Universidade Estadual de Campinas; 2011.

Rizzi L. A restructuring rule. In: *Issues in Italian Syntax*. Dordrecht: Foris; 1982. p. 1-76. [Tradução da versão em italiano: Rizzi L. Ristrutturazione. *Rivista di Grammatica Generativa*. 1976;1:1-54.]

Santiago H. Um estudo do português popular brasileiro em cartas pessoais de “mãos cândidas” do sertão baiano [dissertação]. Feira de Santana: UEFS; 2012.

Schei A. A colocação pronominal na literatura brasileira do século XIX. *Filologia e Linguística Portuguesa*. 2002;5:57-84.

Schei A. A colocação pronominal do português brasileiro: a língua literária contemporânea. São Paulo: Humanitas FFLCH/USP; 2003.

Silverstein M. Hierarchy of features and ergativity. In: Dixon RMW, editor. *Grammatical categories in Australian languages*. Canberra: Australian Institute of Aboriginal Studies; 1976. p. 112-171.

Tarallo F. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: Roberts, I, Kato M, organizadores. *Português Brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp; 1993. p. 35-68.

Torres Morais. MA. Mudança na colocação dos clíticos na história do português. *Boletim da ABRALIN*. 1995;17:32-40.

Vieira SR. Colocação pronominal nas variedades européia, brasileira e moçambicana do português: para a definição da natureza do clítico em Português [tese]. Rio de Janeiro: Universidade Federal do Rio de Janeiro; 2002.