

*A Plataforma PB-Corpus Histórico
e uma investigação da ordem de clíticos
e de sujeitos em jornais brasileiros oitocentistas*

The Platform PB-Historical Corpus and an investigation of the order of clitics and subjects in nineteenth century brazilian newspapers

Marco Antonio Rocha Martins

Universidade Federal de Santa Catarina/CNPq (UFSC)

DOI: <https://doi.org/10.5902/2176148542517>

Resumo: Apresento neste artigo a Plataforma *PB-Corpus Histórico* cujo objetivo é disponibilizar textos em português escritos no Brasil entre os séculos XVIII e XXI. Usando dados extraídos da Plataforma, no quadro teórico da sintaxe diacrônica, analiso a evolução da próclise em ambiente neutro e da ordem do sujeito em textos brasileiros oitocentistas. A hipótese é de que a próclise gerada pela gramática do Português Brasileiro aparece nos textos escritos a partir da segunda metade do século XIX e que nos textos da primeira metade desse século as próclises são geradas por uma gramática do tipo-V2 como o Português Clássico.

Palavras-chave: Clíticos. Sujeitos. *PB-Corpus Histórico*. Sintaxe diacrônica

Abstract: In this article I present the *PB-Corpus Historical Platform* whose objective is to make available texts in Portuguese written in Brazil between the 18th and 21th centuries. In the theoretical framework of diachronic syntax with data extracted from the Platform, I analyze the evolution of proclisis in a neutral environment and the order of the subject in nineteenth-century Brazilian Newspapers. The hypothesis is that the proclisis generated by the BP grammar appears in texts written from the second half of the nineteenth century and that in the texts of the first half of that century the proclises are generated by a V2-type grammar like Classical Portuguese.

Keywords: Clitics. Subjects. *PB-Corpus Historical*. Diachronic syntax

Introdução¹

Trazemos a público, neste artigo, informações sobre a *Plataforma PB-Corpus Histórico*, que constitui um projeto em andamento e tem por objetivo principal armazenar e disponibilizar textos escritos no Brasil no curso dos séculos XVIII a XXI, com programas de interface para coleta e categorização de dados linguísticos. A plataforma está em fase de elaboração e foi carregada, até o momento, com textos oriundos de diferentes *corpora*. Usando dados extraídos da Plataforma, apresentamos resultados de uma investigação da evolução da próclise em ambiente neutro e da ordem do sujeito em textos oitocentistas, buscando argumentos para a hipótese de que a escrita brasileira do século XIX apresenta um quadro de competição entre diferentes gramáticas do português, de modo que, ainda na primeira metade desse século, há resquícios de uma gramática do tipo-V2 como o Português Clássico (PCL). Assim, defendemos a hipótese de que a próclise em ambiente neutro em textos da primeira metade do século XIX é gerada pelo PCL e os padrões gerados pela gramática inovadora do tipo-SV do Português Brasileiro (PB) começam a aparecer em textos escritos a partir da segunda metade desse século. Para além dos padrões gerados pelo PCL e PB, a escrita oitocentista no Brasil é fortemente marcada, ainda, via pressão da norma lusitana, pela gramática do Português Europeu (PE) e esse complexo quadro na escrita brasileira pode ser interpretado como a competição de diferentes gramáticas, seguindo a proposta de Kroch (1989, 2001, 2005).

A hipótese defendida neste artigo encontra respaldo no fato de que, na escrita da primeira metade do século XIX, a estrutura associada à construção [XP]V apresenta indícios de uma gramática que licencia qualquer elemento na posição pré-verbal e não apenas o sujeito, como a gramática do PB ou do PE.

Antes de mais, é importante esclarecer que o conceito de gramática que estamos assumindo é aquele delimitado no quadro teórico da Gramática Gerativa, ou Língua-I, tal como proposto, com muita clareza, em Chomsky (1986). Muito embora, nos valendo da discussão posta em Mendível-Giró (2015), é importante referir que quando trabalhamos com mudança lingüís-

1 Neste artigo, apresento resultados do projeto de pesquisa "Position of the subject and proclisis in neutral contexts [XP]V in 19th century Brazilian writing: Reflexes of a parametric change in BP", financiado por uma bolsa de estudos do Instituto Humboldt/CAPES (Processo número 88881.145464/2017-01) durante o período em que fui professor visitante na Universidade de Colônia/Alemanha. Este projeto está integrado ao projeto "A posição do sujeito pré-verbal e das estruturas [XP-clitic-Verb] na escrita brasileira do século XIX", financiado pelo CNPq com a bolsa de estudos de produtividade PQ-2 (processo 310094/2017-8). Agradeço aos órgãos de fomento e as contribuições dos pareceristas da Revista que, com toda certeza, contribuíram com o resultado final deste artigo.

tica, observada em textos de sincronias passadas (ou textos escritos em temporalidades marcadas pela sócio-história das línguas no curso do tempo), precisamos considerar o conceito de Língua-I relacionado historicamente em distintas gerações de falantes; ou seja, assumimos por gramática um construto teórico denominado de língua/ou gramática histórica, que, grosso modo, pode ser identificado como o somatório do conceito de Língua-I às diferentes temporalidades em distintas gerações de falantes.

O artigo se organiza em quatro seções: em 1. apresentamos um breve panorama de *corpora* históricos do português no Brasil e informações sobre o projeto em curso para a elaboração da *Plataforma PB-Corpus Histórico*; em 2, delimitamos o fenômeno investigado a partir de dados extraídos da Plataforma; em 3, apresentamos os resultados de uma investigação da evolução da próclise em ambiente neutro e a ordem SV na escrita brasileira oitocentista; em 4, concluímos o artigo, sistematizando os argumentos a favor da hipótese defendida de que a escrita no Brasil da primeira metade do século XIX apresenta propriedades de uma gramática do tipo-V2 como o PCl.

1. A *Plataforma PB-Corpus Histórico*: breves notícias

Desde o ressurgimento do interesse pelos estudos em linguística histórica e história do português no Brasil, na década de 1980, que, segundo Castilho (2018), encontram força nos trabalhos pioneiros de Rosa Virgínia Mattos e Silva, na Universidade Federal da Bahia, e de Mary Kato (e, acrescido, de Fernando Tarallo), na UNICAMP, muitos *corpora* voltados à documentação de textos brasileiros têm sido sistematizados. Muitos são os projetos em desenvolvimento e algumas plataformas já foram disponibilizadas na rede mundial. Desses, destacamos os seguintes: o *Projeto para a História do Português Brasileiro* (PHPB) <<https://sites.google.com/site/corporaphpb/>> e os projetos estaduais vinculados ao PHPB nacional, incluindo o *Corpus do Laboratório de História do Português* <<http://www.letras.ufrj.br/laborhistorico/>>, coordenado por Célia Lopes na UFRJ; o *Programa para a História da Língua Portuguesa* (PROHPOR) <<http://www.prohpor.org/#lbit-banco-pb/c1rm>>; o *Corpus Eletrônico de Documentos Históricos do Sertão* (CE-DOHS) <<http://www.tycho.iel.unicamp.br/cedohs/corpora.html>>; e o *Corpus Histórico do português Tycho Brahe* <<http://www.tycho.iel.unicamp.br/corpus/>>².

2 Há outros importantes *corpora* históricos do português aqui não citados porque nos restringimos ao português brasileiro. Para um panorama de outros *corpora* do português consultar Rocha e Marcotulio (2019).

O PHPB, criado e coordenado desde então, até 2019, pelo professor Ataliba de Castilho em 1997, agrega mais de 200 pesquisadores de diferentes IES de muitos Estados brasileiros e tem colocado em pauta uma agenda para a pesquisa em linguística histórica no cenário histórico e sociolinguístico brasileiro. Um dos mais importantes resultados desse importante projeto foi a organização e disponibilização dos *corpora* com textos impressos e manuscritos escritos no Brasil dos séculos XIX e XX com livre acesso³, assim como a organização e publicação de a coleção com doze volumes *História do Português Brasileiro* com base na análise de textos do *corpora* diacrônico compartilhado do Projeto e o trabalho de anos de pesquisa das equipes regionais (CASTILHO, 2018). De acordo com Barbosa (2019, p. 18), a Plataforma PHPB apresenta “uma sistematização de todos os materiais editados pelos membros do Projeto até o Censo de *corpora* do PHPB em junho de 2010”.

É importante destacar a significativa contribuição do PROHPOR, atualmente sob a coordenação de Tania Lobo, que foi fundado em 1990 e permaneceu por muitos anos sob a coordenação da professora Rosa Virgínia Mattos e Silva, na Universidade Federal da Bahia. É um projeto pioneiro sobre documentação histórica para o estudo do português no/do Brasil e tem reunido importantes documentos para descortinar histórias do português, que estão disponíveis no *site* do projeto, assim como tem publicado regularmente obras de referência com análises desses textos. Além do PROHPOR, outro importante projeto desenvolvido em universidades na Bahia, para documentação de textos históricos do português brasileiro, é o CE-DOHS, em Feira de Santana/BA, coordenado por Zenaide Carneiro e Mariana Oliveira.

O *Tycho Brahe*, fruto de um projeto coordenado por Charlotte Galves, na UNICAMP, é um *Corpus histórico* eletrônico anotado com textos em português escritos por autores diversos nascidos entre 1380 e 1881 (cf. informações em <http://www.tycho.iel.unicamp.br/corpus/>). A plataforma desse *corpus* permite realizar buscas de estruturas sintáticas automáticas em grande quantidade de dados nos textos que estão anotados via programa E-Dictor, desenvolvido pelos pesquisadores do projeto: Pablo Faria, Fábio Kepler e Maria Clara Paixão de Sousa.

³ Ver Castilho (2019), e mais especificamente Barbosa (2019), para mais informações sobre o *corpus* diacrônico do projeto PHPB.

De um modo geral, esses *corpora* eletrônicos, com livre acesso, disponibilizam textos em português escritos em diferentes períodos por autores portugueses e brasileiros (quando possível identificar os dados sociais dos escreventes) e estão organizados em gêneros textuais/discursivos variados, tais como cartas particulares e cartas oficiais, documentos notariais, textos de jornais (Editoriais, Anúncios e Cartas de leitores), textos literários (romances, peças de teatro) e outros. Um grande impasse para a pesquisa e coleta de dados linguísticos em número expressivo, na maioria desses *corpora*, é que eles armazenam textos em arquivos em Word ou PDF que podem ser baixados, mas não permitem uma busca rápida de dados específicos, sistematizada por gênero textual/discursivo ou período. O *Corpus Tycho Brahe* e o *CEDOHS* são os únicos, dos aqui citados, que disponibilizam uma versão dos documentos editados em XML que permitem buscas de estruturas sintáticas, mas os demais não apresentam possibilidades de buscas de dados *on-line*, dentro das plataformas, para pesquisa. São disponibilizados os arquivos, organizados em gêneros e regiões, como no caso do PHPB nacional e regionais, que podem ser baixados e as buscas de dados devem ser manuais, com recursos do Word ou PDF.

Com o objetivo de elaborar e disponibilizar uma plataforma interativa, com programas de interface para busca e tratamento de dados de textos escritos em português do/no Brasil, em diferentes sincronias, vimos desenvolvendo, em parceria com Anna Fischer, assistente contratada pela Fundação Humboldt/Alemanha, como projeto resultante do trabalho como *visiting professor* na Universität zu Köln, no período de setembro de 2018 a novembro de 2019, a *Plataforma PB-Corpus Histórico*. O objetivo primeiro dessa Plataforma é armazenar e disponibilizar textos escritos no Brasil no curso dos séculos XVIII a XXI de diferentes *corpora*, de modo que interfaces interativas permitam buscas de dados para estudos de diferentes fenômenos linguísticos, em aplicativos de análises estatísticas, como o “Programa R”, o GoldVARB e outros, por exemplo.

Na organização da Plataforma, tomamos por referência o trabalho que vem sendo desenvolvido no âmbito do PHPB (e reporto à leitura de Barbosa (2019) para uma visão detalhada da constituição do *corpora* desse projeto) e buscamos reunir materiais de diferentes gêneros textuais/discursivos: da esfera dos textos impressos de jornais: **anúncios, cartas de leitores e cartas de redatores/editoriais**; da esfera dos textos manuscritos: **cartas privadas**; e da esfera dos textos

literários: **peças de teatro**. Os textos estão organizados por amostras/ou coleções, para buscarmos preservar as informações das fontes de onde foram extraídos e agrupam textos entre os séculos XVIII a XXI agrupados num período de tempo de cinquenta em cinquenta anos (1701 a 1750; 1751 a 1800; 1801 a 1850; 1851 a 1900; 1901 a 1950; 1951 a 2000; e a partir de 2001). Desse modo, a pesquisa de dados pode ser feita considerando a data de publicação⁴ de cada texto individual, que pertence a uma dada coleção, ou o agrupamento de textos em cada metade de século. A organização conta, ainda, com a distribuição por regiões e Estados e, até o momento, a Plataforma dispõe de textos da região Sul: Rio Grande do Sul, Santa Catarina e Paraná; Sudeste: São Paulo, Rio de Janeiro e Minas Gerais; e Nordeste: Bahia, Pernambuco, Rio Grande do Norte e Ceará.

Os textos depositados foram extraídos em sua maioria dos *corpora* PHPB, aqueles disponíveis em <<https://sites.google.com/site/corporaphpb/>>⁵ e de projetos Estaduais (Rio Grande do Sul, Santa Catarina e Rio Grande do Norte), de acervos do autor deste artigo e de outros *corpora*. Neste momento, a *Plataforma PB-Corpus Histórico* está em fase de elaboração e não está ainda disponível ou hospedada em nenhum domínio.

O diferencial da *Plataforma* em relação aos *corpora* de onde foram extraídos os textos, como dito, reside no fato de possibilitar uma interatividade no manuseamento dos textos e das amostras que permite buscas de dados linguísticos por sistemas de interfaces. Com o uso da *Plataforma*, não é necessário baixar os arquivos para as buscas, que podem ser feitas de forma interativa dentro do sistema. Neste momento, e para testes, foi elaborado e implementado um programa de interface na *Plataforma* que permite a coleta de sentenças com pronomes clíticos, o que possibilita a coleta e categorização para análise de um número robusto de dados.

Como dito, a *Plataforma PB-Cópus Histórico* está em fase de desenvolvimento, de modo que os textos estão sendo preparados e os dados com pronomes clíticos já estão extraídos nesse processo. Nosso objetivo é dar continuidade à elaboração da *Plataforma*, vinculando a sua organização e disponibilização a futuros projetos de parceria entre a Universidade Federal de Santa Catarina/Brasil, via projeto *Portal Línguas e Literaturas/PRINT-CAPES*, a Universidade de Colônia/Alemanha e demais IES parceiras.

4 Para os textos que temos informações sociais do autor, como data de nascimento, por exemplo, elas são disponibilizadas.

5 Acesso em 25 de março de 2020.

A análise da evolução da próclise em ambientes neutros na escrita brasileira oitocentista que apresentamos neste artigo foi feita tendo em vista parte dos dados já coletados via Plataforma PB-Corpus Histórico.

2. O fenômeno: próclise em ambientes neutros ([XP])[XP]V

Seguindo em parte Martins (1994), definimos neste artigo por ambiente neutro ([XP])[XP]V sentenças matrizes, em que o constituinte que antecede o verbo é um sujeito não focalizado, como dado em (1), um sintagma preposicional (2), um advérbio de qualquer natureza (3) ou uma oração subordinada (4):

(1) Verbo precedido de sujeito

[XX,2 CL SC]⁶ **Vocês se lembram** daquela musiquinha que diz assim:
*Choveu, choveu Choveu Canasvieiras encheu Quando chove*R34001

(2) Verbo precedido de sintagma preposicional

[XX,1 A SC] **No armazem de Henrique Schutel vende-se** milho a 1:280 réis o sacco

(3) Verbo precedido de advérbio

[XX,2 CL SC] **Minha filha tomou 18 frascos [de] Peitoral de Cambará e hoje acha-[se]** completamente restabelecida.

(4) Verbo precedido de oração subordinada

[XX,1 CL SC] **Attendendo ao seu longo passado, cumpre nos** todavia levar ao conhecimento de Vossa senhoria que não podemos mais evitar a explosão do nosso operariado que esta se manifestando profundamente desgostoso com a attitude do Senhor Neitsch.

A próclise nesse ambiente, tal como evidenciado num primeiro momento por Martins (1994), é variável e define diferentes períodos na história do português. De modo particular, trata-se de um contexto sintático interessante para a observação da implementação da gramática do PB na escrita brasileira no curso dos séculos XIX e XX, porque marca um divisor de águas entre as gramáticas do portu-

6 Os códigos que precedem os dados codificam o seguinte: Século - XIX.1, XIX.2, XX.1 e XX.2; Gênero textual - LR: cartas de leitores; LJ: cartas de jornalistas; A: anúncio; TP: peça teatral; Estado Brasileiro - SC: Santa Catarina; RJ: Rio de Janeiro; BA: Bahia; PE: Pernambuco; RN: Rio Grande do Norte; CE: Ceará.

guês d'aquém, no Brasil, e d'além mar, em Portugal. Diversos estudos sobre a sintaxe dos clíticos pronominais na história do português escrito no Brasil mostram uma queda no uso da próclise nesse ambiente, da primeira para a segunda metade do século XIX (LOBO 1992; PAGOTTO 1992; SCHEI 2003; DUARTE; PAGOTTO 2005; GALVES; TORRES MORAES; RIBEIRO 2005; CARNEIRO 2005; MARTINS 2009, 2018, a sair; CARNEIRO; GALVES 2010; CAVALCANTE; DUARTE; PAGOTTO 2011). Para a evolução da próclise nesse contexto, observem-se, a título de exemplo, os resultados sistematizados por Martins (a sair), com base em uma amostra de jornais impressos em diferentes estados brasileiros e peças de teatro:

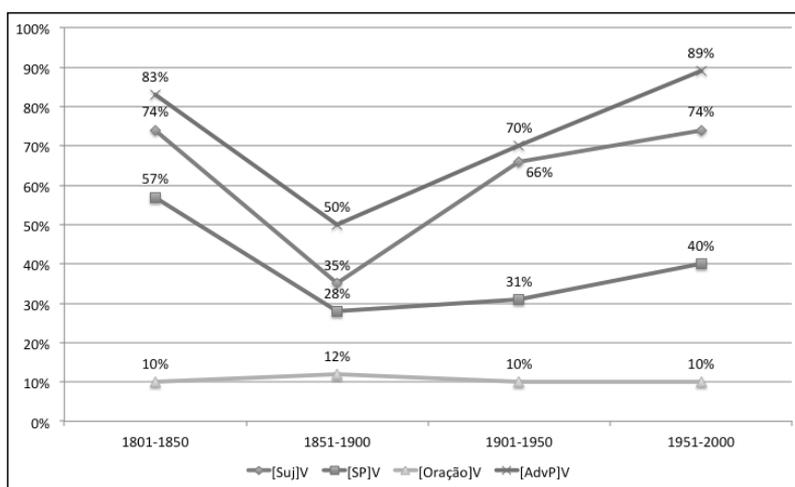


Figura 1 – Próclise em contextos neutros no português escrito no Brasil dos séculos XIX e XX

Há uma queda bastante acentuada da próclise em textos da primeira para a segunda metade do século XIX, exceto em sentenças com orações subordinadas pré-verbais. Essa queda (ou mudança fracassada) na gramática do PB tem sido interpretada como o resultado da pressão da norma lusitana, fixada pelos padrões enclíticos da gramática do PE (cf. PAGOTTO, 1998; CARNEIRO, 2005). Apresentamos aqui outra hipótese para interpretar esses resultados: a próclise gerada pela gramática do PB no contexto ([XP])[XP]V começa a aparecer com significativa expressividade apenas em textos brasileiros oitocentistas da segunda metade do século XIX, de modo que aquela atestada na escrita da primeira metade do século XIX é, ainda, resquício de uma gramática do tipo-V2, como o PCl.

Em teoria gerativa, uma propriedade crucial de uma língua do tipo-V2 é o movimento do verbo para uma posição mais alta na estrutura da sentença, com a obrigatoriedade de um constituinte realizado na posição pré-verbal. Isso não é o que acontece exatamente no PCI, em que a presença desse constituinte em posição pré-verbal é opcional. Seguindo a proposta de Galves e Paixão de Sousa (2017), no entanto, consideramos que o PCI partilha propriedades de línguas do tipo-V2 rígidas como a de não ter uma posição fixa pré-verbal para os sujeitos. Em outras palavras, por língua do tipo-V2 assumimos aquelas que licenciam qualquer elemento, incluindo sujeitos, mas não obrigatoriamente, na posição pré-verbal. Essa é uma propriedade que diferencia o PCI do PB e PE, línguas em que a posição pré-verbal é destinada a sujeitos.

Uma evidência independente para a hipótese que colocamos aqui é o fato de, na escrita brasileira da primeira metade do século XIX, a estrutura associada à construção [XP]V estar associada a de uma gramática que licencia qualquer elemento na posição pré-verbal e não apenas o sujeito, como mostram por exemplo os resultados de Berlinck e Coelho (2018) sobre a evolução da ordem SV e a perda de VS na escrita brasileira.

A hipótese aqui investigada em dados extraídos da Plataforma *PB-Corpus Histórico*, portanto, é de que a escrita brasileira da primeira metade do século XIX reflete, ainda, o padrão proclítico [XP]clV da gramática do PCI e apenas na segunda metade desse século se atesta com mais expressividade a próclise gerada pela gramática inovadora do PB. Nesse sentido e direção, a explicação dada por Pagotto (1998) e Carneiro (2005) para a queda da próclise em contextos neutros na escrita brasileira da segunda metade do século XIX deve ser relativizada. Muito naturalmente, a complexa tessitura sociolinguística brasileira, sobretudo no efervescente cenário no Brasil do século XIX, traz um complicador a mais a esse quadro de variação, que aqui interpretamos como competição de diferentes gramáticas do português: há uma forte imposição da norma enclítica lusitana. Esse quadro, de um alto índice de ênclises nos textos em solos brasileiros nesse período, é inegavelmente atribuído à pressão da norma lusitana nos textos escritos no Brasil. O que quero destacar é que, nesse quadro, não há uma queda de próclise na gramática do português brasileiro em direção à ênclise nesse contexto, ou uma mudança fracassada, mas que a próclise do PB nesse contexto começa a aparecer nos textos da segunda metade desse século e que aquelas da primeira metade tem fortes indicadores de estarem associadas à gra-

mática do PCl. Mas é inegável que a pressão da norma lusitana impõe seu padrão enclítico, identificada principalmente nos textos da segunda metade dos séculos XIX.

3. A análise: ordem de clíticos em ambiente neutro ([XP])[XP]V e de sujeitos em jornais oitocentistas

Com o propósito de investigar a correlação entre a posição de clíticos em ambiente neutro ([XP])[XP]V, conforme definido na seção 2, e a ordem SV em sentenças matrizes na escrita brasileira oitocentista, apresento uma análise de regra variável de 533 dados retirados de cartas de leitores, cartas de redatores e anúncios de jornais do século XIX em textos da Plataforma *PB-Corpus Histórico*. A análise contempla dados de textos de cinco estados brasileiros: Santa Catarina, Rio de Janeiro, Bahia, Pernambuco e Ceará. Foram consideradas ocorrências com clíticos em ambiente neutro ([XP])[XP]V, não ativador de próclise, em que o constituinte que antecede imediatamente o verbo é um sujeito, um PP ou um advérbio, conforme dados em (1)-(3) acima. Foi realizada uma análise multivariada com os programas do pacote estatístico GoldVarb (2005), com a aplicação da variável dependente “colocação do clítico” correlacionada aos seguintes fatores: (a) posição superficial do verbo, em segunda [XP]V ou terceira posição [XP][XP]V); (b) natureza do constituinte que antecede imediatamente o verbo em contextos ([XP])[XP]V; (c) ordem e realização do sujeito; (d) natureza do sujeito expresso – pronominal, DP, Complexo = DP + relativa, oracional); (e) forma do clítico; (f) gênero do texto; e (g) estado onde fora publicado o texto.

A evolução em termos percentuais da próclise em sentenças matrizes com sujeitos pré-verbais e da ordem SV na amostra foi a seguinte:

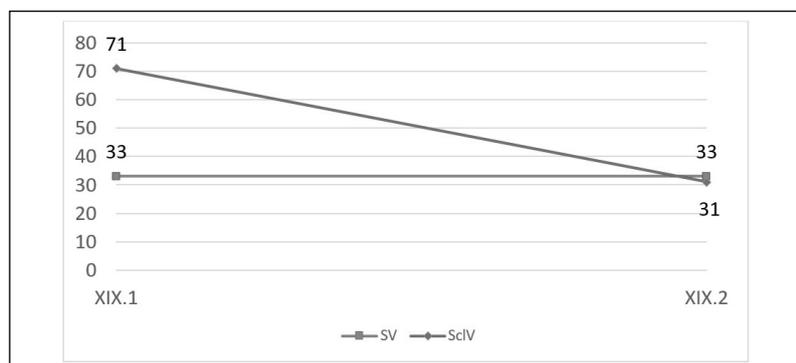


Figura 2 – Próclise em ambiente neutro ([XP])[DP]V e SV em textos de jornais brasileiros oitocentistas

A evolução na próclise no contexto [Sujeito]clíticoVerbo segue aquela encontrada em estudos anteriores e sistematizada na Figura 1. Numa análise particularizada da escrita oitocentista desta análise, destaco que, em textos da primeira metade do século XIX, há uma elevada taxa de próclise em sentenças matriz com sujeito pré-verbal (71%) e a ordem SV é de 33%; em textos da segunda metade desse século, há uma abrupta queda da próclise em SV (que cai para 31%) e a taxa de SV se mantém em 33%⁷. Esse quadro parece sustentar a hipótese de que as próclises geradas em ambiente neutro ([XP])SV em textos da primeira metade do século XIX não estão associadas a uma gramática do tipo-SV com anteposição do sujeito, como o PB, conforme proposta de Martins (2009) e Martins, Cavalcante e Coelho (2020). Como mostram os resultados desses últimos autores, em textos brasileiros da primeira metade do século XIX, há indícios de uma gramática do tipo-V2 com sujeitos nulos com próclise em contextos ([XP])[DP]V; e, em textos a partir da segunda metade desse século, evidenciam-se propriedades de uma gramática do tipo-SV, com sujeitos lexicalizados e próclise em ambientes ([XP])[DP]V (conforme, ainda, resultados de BERLINCK, 1988, 1989, 1995; COELHO, 2006; CAVALCANTE, 2018; BERLINK; COELHO, 2018).

Na análise multivariada com as variáveis acima elencadas, o programa selecionou como estatisticamente relevante, na melhor rodada com *Log likelihood* de -303.954 e *Significance* de 0.000, as seguintes variáveis nessa ordem de relevância: (1) natureza do sujeito expresso, (2) natureza do constituinte que antecede imediatamente o verbo em contextos ([XP])[XP]V, (3) forma do clítico e (4) ordem e realização do sujeito. Apresento os resultados dessas variáveis no que segue, começando pela natureza do sujeito expresso.

7 Como mostram Coelho e Berlinck (2018), é na escrita brasileira do século XX que se observa (ou que deixa mostrar a sua tinta com mais expressividade, para retomar Tarallo, 1993) a gramática do PB em relação à rigidez de uma ordem SV(O). Nas palavras das autoras, “no século XIX, VS manifesta resquício da sintaxe do PA e do PC no que se refere aos padrões de inversão germânica [XP]VS e românica V[XP]S, especificamente com verbos transitivos” (p. 378). Esse quadro justifica que nos dois períodos da escrita oitocentista aqui analisados encontremos, ainda, baixas e similares frequências de SV.

Tabela 1 – Próclise em ambiente neutro ([XP])[DP]V, por natureza do sujeito expresso

	Occ. - %	Prob.
Pronome Pessoal	38/45 - 84,4%	0,92
Demonstrativo	6/8 - 75%	0,86
Complexo	10/25 - 40%	0,58
DP	73/199 - 36,7%	0,55
Oracional	8/104 - 7,7%	0,15
TOTAL	187/378 - 49,5%	

Marco Antonio
Rocha Martins

190

Sujeitos pronominais pessoais e demonstrativos condicionam a próclise com pesos relativos de 0,92 e 0,86 em oposição a sujeitos complexos, DP e oracionais, com pesos de 0,58, 0,55 e 0,15. Martins (2009, 2018) mostrou já que pronomes pessoais são condicionadores de próclise na história do português escrito no Brasil e os resultados da escrita oitocentista confirmam esse resultado. Os sujeitos oracionais desfavorecem a próclise e muito provavelmente o motivo seja o elevado número de construções formulaicas em anúncios oitocentistas com relativas livres encabeçadas pelo pronome relativo “quem”, conforme (5).

(5) [XIX,1 A RJ] NOTICIAS MARITIMAS PARA LONDRES Para passageiros sómente. O superior Bergantim Inglez GEORGE e WILLIAN, Capitão GEORGE NICKLESON. Forrado e pregado de cobre. A [espaço] tem superiores commodos para passageiros. **Quem quizer hir de passagem dirija-se** aos Consignatorios Philips e *Companhia* ou aos Corretores de Navios hudson e Weguelin. Rua Direita número136.

A segunda variável selecionada foi a natureza do constituinte que antecede imediatamente o verbo em contextos ([XP])[XP]V e os resultados estão na tabela 2:

Tabela 2 – Próclise por natureza de XP que antecede imediatamente o verbo no contexto ([XP])[XP]V

	Occ. - %	Prob.
Advérbio	64/85 - 75,3%	0,79
PP	51/126 - 40,5%	0,44
Sujeito	72/67 - 43,1%	0,37
TOTAL	187/378 - 49,5%	

Advérbios (peso relativo de 0,79) em oposição a Sintagmas Preposicionais (0,44) e sujeitos (0,37) imediatamente antecedentes ao verbo condicionam a próclise. Esse resultado era já esperado, pois incluímos aqui todos os advérbios por considerar que estudos anteriores sobre o português escrito no Brasil do século XIX mostraram variação na posição de clítico em sentenças com todos os tipos de advérbios, inclusive os modais (MARTINS, 2009, 2012, 2018; CARNEIRO 2005), conforme dados em (6)

- (6) ênclise com advérbios ativadores de próclise
- a. [XX.2 LR RJ] **Já deixei-A** na Motocar e o mecânico me disse que era “azinhavre”, e que “produto nacional era uma droga mesmo”!
 - b. [XIX.1 LR BA] [...] e **assim comprei-A** muito bem
 - c. [XIX.2 LR BA] Em conclusão declaro ao Senhor Philantropo que o Senhor José Revault é incapaz de contrariar a liberdade das ditas crias, pois **sempre dizia-ME** que ellas erão forras.
 - d. [XIX.2 A BA] [...] e **tambem vendem-SE** folhas avulsas a 60 réis.

Martins, Cavalcante e Coelho (2020), numa análise que considerou rodadas estatísticas distintas com dados da primeira metade do século XIX e dados da segunda metade do século XIX e do século XX, mostram que os SPs perdem força no condicionamento da próclise em oposição a sujeitos em textos oitocentistas da segunda metade do século. Essa mudança na força de condicionamento da próclise em ambientes neutros de PPs pode ser interpretada como um enfraquecimento ou mudança de uma gramática do tipo-V2 para uma gramática do tipo-SV, como aquela que marca a passagem do PCI para o PB e PE (GALVES; PAIXÃO DE SOUSA, 2017).

O tipo de clítico foi a terceira variável selecionada e os resultados mostram que há condicionamento quando aparecem dois clíticos, contratos ou não, e do clítico *te*, com pesos relativos de 0,87 e 0,84 em oposição aos demais⁸:

8 Encaminho a Martins (2018) para uma possível explicação desse condicionamento.

Tabela 3 – Próclise por forma do clítico no contexto ([XP])[XP]V

	Occ. - %	Prob.
Dois clíticos	4/6 – 66,7%	0,87
TE	8/10 – 80%	0,84
O/A	27/45 – 60 %	0,65
LHE	24/41 – 58,5%	0,64
NOS	11/20 – 55%	0,61
SE	101/338 – 29,9%	0,45
ME	22/69 – 31,9%	0,40
VOS	2/2 – 50%	0,22
Total	199/533 – 37,3%	

Marco Antonio
Rocha Martins

192

A realização e posição do sujeito também se mostrou uma variável relevante para o condicionamento da próclise, conforme dados na tabela a seguir:

Tabela 4 – Próclise por realização e posição do sujeito

	Occ. - %	Prob.
SV	93/284 – 32,7%	0,59
VS	42/96 – 43,8%	0,44
nulo	33/88 – 37,5%	0,26
TOTAL	168/468 – 35,9%	

Esse resultado é muito interessante porque mostra que o que está em jogo na escrita brasileira oitocentista parece ser a oposição entre sujeitos preenchidos *versus* sujeitos nulos: sujeitos preenchidos antepostos (0,59) ou pospostos (0,44) são condicionadores da próclise no contexto ([XP])[DP]V em oposição a nulos (0,26). Nas construções com sujeitos nulos, em que há uma forte inclinação para o uso da ênclise, encontramos na posição pré-verbal advérbios, PPs ou orações subordinadas, conforme ilustram os dados em (7).

(7) Vcl com sujeitos nulos

- a. [XIX,1 CL BA] calumnioso, porque todos sabem, que eu não sou capaz de fazer contractos simulados; e injusto, porque, quando comprei á dita venda, o vendedor Francisco de Moraes, não tinha impedimento algum: e **assim comprei-a** muito bem
- b. [XIX,1 CL PE] **Em 1830 tornou-se** original em suas votações, e ultimamente aceitou a Pasta da Fazenda.

Esse quadro traz evidências a favor da hipótese que aventamos neste artigo de que há uma correlação entre a evolução da próclise no ambiente neutro ([XP])[XP]V e da ordem SV do sujeito gramatical na escrita brasileira oitocentista. Interpretamos esse quadro como o reflexo da competição de diferentes gramáticas na derivação da próclise, de modo que, por um lado, apesar de superficialmente similar, a próclise no contexto [XP]clV em textos da primeira metade do século XIX é gerada por uma gramática conservadora com sujeitos nulos e com propriedades de uma língua do tipo-V2, conforme (8); de outro lado, em textos da segunda metade do século XIX (e do século XX, como mostram Martins, Cavalcante e Coelho, em preparação) atuam forças de uma gramática do tipo-SV, com sujeitos preenchidos e antepostos em SV, como o PB, como em (9) e (10). A atuação da gramática do PB realinha a colocação do pronome clítico em próclise de modo que o que se observa no curso da escrita oitocentista no Brasil não é uma queda da próclise no ambiente neutro ([XP])[XP]V, mas a implementação da próclise gerada pela gramática do PB em textos escritos a partir da segunda metade do século XIX.

(8) [XP]V

[XIX,1 CR RJ]

Ilustríssimo e Excelentíssimo Senhor **Pela segunda vez nos dirigimos** a Vossa Excelência para de novo rogarmos-lhe a publicação do parecer da comissão encarregada de examinar os excandalosos roubos do arsenal de guerra, o que já Vossa Excelência de a muito que o deveria ter feito para poder desmentir os boatos que por ahi se tem espalhado de que Vossa Excelência e o seu Official maior são os unicos criminosos.

(9) [S]V

[XX,1 CL PE] **O Papa nos diz:** “A paz não será autentica senão quando ela for o fruto da justiça.

(10) [XP]SV

a. [XX,2 CL CE] Após o governo democrático e de processo de Juscelino, com João Goulart na presidencia, uma onda de subversão se espalhou por todo o país, comandada por uma minoria sedenta de reformas, cega de paixões, que ameaçava incendiar a nação se, suas teses não fossem postas em prática

b. [20,2 CR CE] Para nós, do Nordeste, que vivemos, todos os anos, o “suspense” das chuvas, sem dispor de dados por mais elementares que sejam, para orientar a agricultura e a pecuária, as declarações do técnico da OMM se revestem de inegável significação.

É importante esclarecer que a empiria e os resultados estatísticos de correlação e peso relativo parecem mostrar que a ausência de uma rigidez SV e a próclise no contexto ([XP])[XP]V nos textos da primeira e segunda metade do século XIX se relacionam, mas tal correlação superficializa diferentes caminhos de mudanças na gramática do PC em direção ao PB: a perda das propriedades de uma língua do tipo-V2 e a derivação da próclise no contexto ([XP])[XP]V.

O PCl possui propriedades de uma língua do tipo-V2, com frontamento de qualquer constituinte, incluindo, mas não obrigatoriamente, o sujeito; há a derivação da próclise no contexto ([XP])[XP]V e uma explicação para a próclise nesse contexto pode ser encontrada em Galves, Torres Morais e Ribeiro (GTMR, 2005), que propõem que os clíticos são afixados à categoria funcional *Inf* e uma restrição gramatical *não inicial* que se no domínio do primeiro X-barra da estrutura está ativa.

O PB perdeu as propriedades de uma língua do tipo-V2, passando a SV irrestrita; há a derivação da próclise no contexto ([XP])[XP]V, mas em diferentes contextos estruturais do PCl porque, seguindo GTMR (e ver, ainda, Martins, 2018), os clíticos são afixados a categoria lexical verbal (num nível mais baixo na estrutura da sentença) e uma restrição gramatical *não inicial* não está mais ativa em qualquer domínio⁹.

4. Para concluir, evidências independentes para o reflexo do PCl na escrita brasileira oitocentista

Com a análise de jornais brasileiros oitocentistas, numa amostra de dados extraídos da plataforma *PB-Corpus Histórico*, apresentamos neste artigo evidências empíricas que justificam uma correlação entre a evolução de dois fenômenos sobejamente descritos na história do português escrito no Brasil: a próclise em ambiente neutro [(XP)][XP]V e enrijecimento da ordem Sujeito-Verbo, em estruturas [XP]V com sujeito pré-verbal (como (11a.) ou em estruturas [(XP)][XP]V também com sujeito pré-verbal (como em (11b.).

⁹ Um movimento distinto pode ser encontrado no PE em que há a perda das propriedades de uma língua do tipo-V2, mas a ênclise é derivada no contexto [(XP)][XP]V. Uma explicação para essa diferença pode ser encontrada em Galves, Britto e Paixão de Sousa (2005).

(11) ([XP])[Sujeito]clV

a. [20,2 CL PE] **Wamberto se insere** no rol dos meus mais caros e achegados amigos.

b. [20,1 CL PE] No Carnaval **ele se diverte**, na Páscoa penas no Cristo morto.

Sistematizamos argumentos a favor da hipótese de que a próclise gerada pela gramática do PB no ambiente neutro ([XP])[XP]V aparece com mais expressividade apenas em textos brasileiros oitocentistas da segunda metade do século XIX, momento em que se revelam expressivamente propriedades de uma gramática do tipo-SV, em oposição a uma gramática do tipo-V2.

Outra evidência independente para a hipótese é que em textos brasileiros da primeira metade do século XIX são encontradas vestígios de uma gramática do tipo-V2 como dados com ordem VS irrestrita (conforme (12), com interpolação de constituintes diferentes do marcador de negação frásico “não” (13) e com subida de clíticos em sentenças matrizes não passivas (14).

(12) [XP]VS, em contextos transitivos (BERLINCK; COELHO, 2018)

a. [XIX,1 CP RJ] *Em 2 do corrente escrevi-te dando conta das despesas que fez o **Chico** durante o mez passado.*

b. [XIX,1 A RJ] *Domingo 21 do corrente deve ter lugar a festa da padroeira, prega ao evangelho o **Padre José Herculano da Costa Brito**, (...)*

(13) Interpolação de “não” e outros constituintes

a. [XIX,1 CL RJ] *Eu **ME não julgo** criminoso e quando for chamado a contas apresentarei os titulos de minha defeza se he que os innocentes se justificação.*

b. [XIX,1 CR RJ] *Não nos enganamos; a espionagem que desde logo se desenvolveo em grande escalla, **NOS não deixou** mais; nossa casa, nossos amigos erão seguidos por toda a parte, e os proprietarios das diversas typographias a que nos dirigimos, forão ameaçados com deportações, visto alguns serem estrangeiros!*

c. [XIX,1 CL BA] *Implantem-se no animo nacional | novos habitos, desperte-se, cultive-se e estenda se a iniciativa individual, e a organização admi | nistrativa, ainda permanecendo a mesma, justificar-se-ha da imputação que **lhe hoje fazem**.*

d. [XIX,1 CL CE] **ME muito afoítesa**, senhor redactor pedirem se taes attestados: os dois primeiros senhores ja lhe diceraõ abertamente, que não lhe davaõ attestados

(14) Subida de clítico em construções não passivas

a. [XIX,1 CR CE] O objetivo de tal união **SE pôde precisar** sem dificuldades.

b. [XIX,1 A CE] O papel **NOS deverão ser fornecido**, e a[s] out[r]as impre[ss]oe[s] serão feitas, segundo o que se convencionar.

Marco Antonio
Rocha Martins

196

REFERÊNCIAS

BARBOSA, A. G. A Plataforma de *corpora* do PHPB: uma apresentação ad infinitum. In: CASTILHO, A, T de. **História do Português Brasileiro** – corpus diacrônico do português brasileiro. São Paulo: Contexto, 2019, pp. 16-67.

BERLINCK, R. de A. A construção V SN no português do Brasil: uma visão diacrônica do fenômeno da ordem. In: TARALLO, F. (org.). **Fotografias sociolinguísticas**. São Paulo: Pontes, 1989, p. 95-112.

BERLINCK, R. de A.; COELHO, I. L. A ordem do sujeito em construções declarativas na história do português brasileiro. In: CYRINO, S.; TORRES MORAIS, M. A. (Orgs.) **Mudança sintática do português brasileiro: perspectiva gerativista**. São Paulo: Contexto, 2018, p. 308-381.

CASTILHO, A, T de. **História do Português Brasileiro** – corpus diacrônico do português brasileiro. São Paulo: Contexto, 2019.

CARNEIRO, Z. **Cartas brasileiras (1809-1904): um estudo linguístico-filológico**. Tese de Doutorado. Universidade Estadual de Campinas, 2005.

CARNEIRO, Z.; GALVES, C. Variação e gramática: colocação de clíticos na história do português brasileiro. **Revista de Estudos Linguísticos**, Belo Horizonte, 18. 2: 7-38, 2010.

CAVALCANTE, S. R. de O.; DUARTE, M. E. L.; PAGOTTO, E. Clíticos nos Século XIX: uma questão de posição social?. In: Dinah Callou; Afranio Barbosa. (Org.). **A Norma Brasileira em construção: cartas a Rui Barbosa (1866 A 1899)**. Rio de Janeiro: Fundação Casa de Rui Barbosa / UFRJ, 2011, v. 1, p. 167-217.

CAVALCANTE, S. R. de O. Revisitando as construções com se na história do português brasileiro. In: CYRINO, S.; TORRES MORAIS, M. A. (Orgs.) **Mudança sintática do português brasileiro: perspectiva gerativista**. São Paulo: Contexto, 2018, p. 382-419.

CAVALCANTE, S. R. de O. Mudança na posição do sujeito em cartas pessoais brasileiras: a ordem VS e o estatuto informacional do sujeito. **DIADORIM (RIO DE JANEIRO)**, v. 20, p. 101-121, 2018.

COELHO, I. L. Variação na sintaxe: estudo da ordem do sujeito no PB. In: RAMOS, J. (org.). **Estudos sociolinguísticos: quatro vértices do GT da ANPOLL**. Belo Horizonte: FALE: Ed. da UFMG, 2006, p. 84-99.

COELHO, I. L.; MARTINS, M. A. A diacronia em construções XV na escrita catarinense. **Fórum Lingüístico**, Florianópolis, v.6, n.1, p. 73-90, jan-jun, 2009.

CHOMSKY, N. **O conhecimento da Língua** – sua natureza, origem e uso [1986]. [tradução de Anabela Gonçalves e Ana Teresa Alves]. Lisboa: Caminho, 1994 [p. 21-71]

DUARTE, M. E. L.; PAGOTTO, E. Gênero e norma: avós e netos, classes e clíticos no final do século XIX. In: Lopes, Célia Regina dos Santos. (Org.). **A Norma Brasileira em Construção: fatos lingüísticos em cartas pessoais do século 19**. 1a.ed.Rio de Janeiro: In-Fólio, 2005, p. 67-82.

GALVES, C; PAIXÃO DE SOUSA, M C. The change in the position of the verb in the history of Portuguese: Subject realization, clitic placement and prosody. **Language**, v. 93, 152-180, 2017.

GALVES, C, NAMIUTI, C, PAIXÃO DE SOUSA, M C. Novas perguntas para antigas questões: a periodização do português revisitada. In Endruschat; Kemmler; Schäfer-Prie. **Grammatische Strukturen des Europäischen Portugiesisch**.Tubingen: Calapinus Verlag, 45-75, 2006.

GALVES, C. M. C.; RIBEIRO, I. M. de O.; MORAES, M. A. T. Syntax and morphology in the placement of clitics in European and Brazilian Portuguese. **Journal of Portuguese Linguistics**, Lisboa, v. 4, n.2, p. 143-177, 2005.

GALVES, C.; BRITTO, H.; PAIXÃO DE SOUSA, M. C. The Change in Clitic placement from Classical to Modern European Portuguese: Results from the Tycho Brahe Corpus. **Journal of Portuguese Linguistics**, Vol. 4, n.1, Special Issue on variation and change in the Iberian languages: the Peninsula and beyond, pp. 39-67, 2005.

KROCH, A. Reflexes of Grammar in Patterns of Language Change. **Language Variations and Change**, 1, 199-244, 1989.

LOBO, T. **A colocação dos clíticos em Português. Duas sincronias em confronto**. Dissertação de Mestrado. Universidade de Lisboa, Faculdade de Letras, 1992.

MARTINS, A. M. Posição dos pronomes pessoais clíticos. In: RAPOSO, E. B. P. *et al.*, **A Gramática do Português**. Vol II. Lisboa: Fundação Calouste Gulbenkian, 2013, p. 2231-2302.

MARTINS, A. M. **Clíticos na história do português**. Tese de doutoramento, Universidade de Lisboa, Faculdade de Letras, 1994.

MARTINS, M. A.; CAVALCANTE, S. R de O.; COELHO, I. L. Ordem do sujeito e colocação de clíticos: reflexos de competição de gramáticas do português na escrita brasileira dos séculos XIX e XX. **Cadernos de Estudos Linguísticos** (UNICAMP), v. 62, p. 1-22, 2020.

MARTINS, M. A. **Competição de gramáticas do português na escrita catarinense dos séculos 19 e 20**. 2009. 326p. Tese (Doutorado em Linguística) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2009.

MARTINS, M. A. A sintaxe dos pronomes pessoais clíticos na história do português brasileiro. In: CYRINO, S.; TORRES MORAIS, M. A. (Orgs.) **Mudança sintática do português brasileiro: perspectiva gerativista**. São Paulo: Contexto, 2018, p. 150-209.

MARTINS, M. A. Micro-variation and parametric change: the proclisis in ‘neutral contexts [XP] V’ in Brazilian writing. Berlin: **PhiN Philologie**, a sair.

MENDÍVIL-GIRÓ, J. L. **El cambio lingüístico**. Sus causas, mecanismos y consecuencias. Madrid: Síntesis, 2015.

PAGOTTO, E. 1992. **A posição dos clíticos em português: um estudo diacrônico**. Dissertação de Mestrado, Instituto de Estudos da Linguagem da Universidade Estadual de Campinas: Campinas.

PAGOTTO, E. 1998. Norma e condescendência, ciência e pureza. **Língua e instrumentos lingüísticos**, vol. 2, 49-68.

TARALLO, F. Diagnosticando uma gramática brasileira: o português d’aquém e d’além mar ao final do século XIX. In. Kato e Roberts. **O português brasileiro: uma viagem diacrônica**. Campinas, SP: Unicamp. 1993, p. 69-105.

*A Plataforma
PB-Corpus
Histórico e uma
investigação
da ordem
de clíticos e
de sujeitos
em jornais
brasileiros
oitocentistas*

199

