

A competição de gramática(s) em José de Alencar

Grammar(s) competition in José de Alencar

Williane Silva Corôa¹

Paulo Ângelo de Araújo-Adriano²

Resumo: Observando diferentes gêneros textuais escritos por diferentes autores, alguns trabalhos em diacronia vêm sugerindo que tal diferença influencia os resultados da pesquisa (ver GALVES, 2018; BIAZOLLI; BERLINCK, 2021; ARAÚJO-ADRIANO; CORÔA, 2022). Neste estudo, abordamos a influência do gênero textual nos resultados diacrônicos, argumentando que diferentes gêneros textuais podem moldar a gramática utilizada pelos autores, refletindo diferenças entre a língua escrita e a língua falada, em um processo de Competição de Gramáticas (KROCH, 1989). O objetivo foi investigar a relação entre o gênero textual e a gramática internalizada de um mesmo autor. Para isso, foram analisadas a expressão do Tempo presente (forma analítica *estar* + gerúndio *versus* forma sintética) e a posição do clítico com o verbo em primeira posição (próclise *versus* ênclise) em dois textos de um mesmo autor, José de Alencar, um romance e um texto teatral, a fim de verificar se os gêneros influenciam o uso gramatical. Os dados foram coletados com o auxílio da ferramenta *CorpusSearch* (RANDAL ET AL., 2009) para o texto com anotação morfológica e manualmente para o texto sem anotação. Os resultados preliminares confirmam em parte a hipótese levantada. Tanto para a expressão do Tempo presente quanto para a posição do clítico com o verbo em primeira posição vê-se manifestado o processo de competição de gramáticas, pois Alencar utiliza ora a gramática nuclear ora a gramática periférica. Em suma, o estudo buscou argumentos que apoiem a tese de que as formas linguísticas utilizadas pelos falantes ao longo do tempo são influenciadas pelos gêneros textuais, mostrando a existência de variação diagênica.

Palavras-chave: Competição de gramática(s). Gênero Textual. Gramática nuclear. Gramática periférica. Variação diagênica.

Abstract: Observing different textual genres written by different authors, some diachronic studies have suggested that such differences influence the research results (see GALVES, 2018; BIAZOLLI; BERLINCK, 2021; ARAÚJO-ADRIANO; CORÔA, 2022). In this paper, we address the influence of textual genre on diachronic results, arguing that different textual genres may shape the grammar used by authors, reflecting differences between written and spoken language. The aim was to investigate the relationship between textual genre and the internalized grammar of the same author. To do so, we analyzed the expression of the present Tense (analytical form *estar* + gerund *versus* synthetic form) and the position of the clitic when the verb is in the first position (proclisis *versus* enclisis) in two texts by the same author, José de Alencar, a novel and a comedy play, to verify if genres influence grammatical usage. The data were collected with the assistance of the *CorpusSearch* tool (RANDAL ET AL., 2009) for the text with morphological annotation and manually for the text without annotation. The preliminary results partially confirm the hypothesis raised. Both for the expression of the present tense and for the position of the clitic with the verb in the first position, the process of grammar competition is observed, as Alencar uses either the nuclear grammar or the peripheral grammar. In summary, the study seeks evidence supporting the thesis that the linguistic forms used by speakers over time are influenced by textual genres, showing the existence of diagenic variation.

Keywords: Grammar(s) competition. Textual genre. Core grammar. Peripheral grammar. Diagenic variation.

1 INTRODUÇÃO

Dentro das tarefas do linguista histórico³, mais especificamente do linguista diacrônico, diferentes decisões precisam ser tomadas, seja em relação ao recorte histórico, seja em relação

¹ Universidade do Estado da Bahia. E-mail: wcoroa@uneb.br.

² USP. E-mail: pauloangelo@usp.br

³ Distinguimos, neste texto, linguista histórico do linguista diacrônico no que se refere à mudança linguística como objeto de investigação. A despeito do seu sentido *lato sensu* que inclui análise de dados em (diferentes) períodos da língua, no sentido *stricto sensu* mudanças nas línguas podem receber o olhar da linguística histórica e

ao modo como a gramática de falantes de outros períodos vai ser representada. Nesse aspecto, estão envolvidas questões sobre a data adotada para a periodização⁴ – data de publicação do texto ou data de nascimento do autor – e o gênero textual a ser analisado.

Em relação à primeira questão, Araújo-Adriano e Corôa (2022) argumentam que uma análise diacrônica baseada apenas na geração biológica, isto é, data de nascimento do autor, não captura possíveis variações na gramática do mesmo autor em períodos distintos. Investigando a variação entre *estar + a + infinitivo* e *estar + gerúndio*, estruturas progressivas prototípicas na variedade europeia e brasileira, respectivamente, em peças de teatro brasileiras, dentre as quais duas são escritas por Martins Pena em 1833 e 1845, os autores mostram que na primeira peça o escritor é muito mais conservador, fazendo uso da variante lusitana em demasiado (33%), enquanto na sua segunda peça a influência europeia é consideravelmente menor (9%), prevalecendo a variante prototípica brasileira. Justificada, discutivelmente, pela pressão portuguesa no começo do século 19, essa diferença não poderia ser atestada, caso fosse considerada apenas a data de nascimento de Martins Pena (1815) para a computação do fenômeno (ver também CARNEIRO, 2005; GALVES; BRITO; PAIXÃO DE SOUSA, 2005).

Já em relação à segunda questão, a do gênero textual, algumas pesquisas sugerem que gêneros textuais diferentes influenciam os resultados diacrônicos (ver também ROMAINE, 1982; GALVES, 2006). Por exemplo, ao ponderar os capítulos em Cyrino e Torres Morais (2018), em que os autores utilizaram como base de dados matérias de jornais, Galves (2018) conclui que a comparação dos resultados apresentados ali com resultados de outros estudos, baseados em obras teatrais, gera certo paradoxo: em um primeiro momento, a incompatibilidade aparece como complicador, já que se coloca uma questão sobre a representatividade dos dados, em um segundo momento, diferentes gêneros textuais aparecem como condicionadores relevante que moldam a “língua escrita e distorcem a imagem da evolução da língua no seu uso coloquial” (GALVES, 2018, p. 442).

Na mesma direção, Araújo-Adriano e Corôa (2022) argumentam que a escolha do gênero textual não deve ser feita de forma ingênua, pois a decisão sobre o gênero impacta diretamente os resultados. Contrastando peças de teatro e cartas e atas, os autores mostram que enquanto as estruturas progressivas brasileiras são insensíveis ao gênero textual, o tipo de texto condiciona a colocação clítica: a despeito de serem mais próximas à oralidade, as peças tendem a ser mais conservadoras, diferentemente das cartas e atas, majoritariamente proclíticas. O que esses trabalhos sugerem é que autores, e em última instância falantes, ativam

diacrônica (MATTOS E SILVA, 2008). Enquanto a primeira considera aspectos extralinguísticos/sociais conjugados a aspectos intralinguísticos, tendo como precursores a sócio-história (ROMAINE, 1982) e a teoria da variação e mudança (WEINREICH; LABOV; HERZOG, 1968), a segunda considera, *grosso modo*, aspectos intralinguísticos, baseados em proponentes como Lightfoot (1979) em um modelo gerativista, ainda que recentemente pesquisadores inseridos nesse modelo têm argumentado que aspectos sociais são cruciais para o melhor entendimento da mudança (ver PINTO; ANDRADE, 2019). É nesse sentido que distinguimos linguista histórico de linguista diacrônico: aquele conjuga aspectos extralinguísticos enquanto este, *grosso modo*, inclui apenas aspectos intralinguísticos.

⁴ Apesar de ser extremamente importante, este critério não será analisado neste artigo. Para mais informações, ver também Araújo-Adriano; Corôa (2022).

gramáticas diferentes, moldando-as a depender do gênero na qual a gramática é veiculada, independente da gramática internalizada durante o processo de aquisição da linguagem.

Durante esse processo, o falante internaliza diferentes regras da língua às quais está exposto, formando sua própria gramática nuclear (CHOMSKY, 1986; KATO, 2013). Essa gramática é adquirida de maneira natural, espontânea e sem correções, diferentemente da gramática periférica, uma gramática formada por empréstimos, fósseis linguísticos inseridos *via* escolarização (KATO 1999, 2013; KATO; CYRINO; CORREA, 2009; ARAÚJO-ADRIANO, 2020). Esse é o caso das formas de futuridade em português brasileiro. Araújo-Adriano (2020) mostra que enquanto crianças brasileiras adquirem de forma natural a forma analítica de futuro, *viz.* *ir* + infinitivo, a forma sintética, *e.g.*, ‘quererei’, presente no PB até o fim do século XIX, só é adquirida ao longo da escolarização, tendo seu uso aumentado à medida que o nível de escolaridade também aumenta. O interessante é que o uso de forma sintética só aparece na escrita, mas não na fala. A gramática internalizada do falante seria, pois, resultado de uma competição entre duas gramáticas, quais sejam: as gramáticas nuclear e periférica (KROCH, 1989).

No âmbito do *labor* do linguista diacrônico em que decisões metodológicas são cruciais e considerando que em Araújo-Adriano e Corôa (2022) foram discutidas as decisões sobre a oposição data de nascimento e data de publicação no mesmo autor, *viz.* Martins Pena, neste trabalho, objetivamos discutir a relação entre o gênero textual e a gramática internalizada de um mesmo autor. Para tanto, investigamos dois textos de José de Alencar – um romance, *Iracema* (1865), e um texto teatral, *O demônio familiar* (1857) –, ambos disponíveis no *Corpus Histórico do Português Tycho Brahe*, parte do *Corpus Tycho Brahe Brasil* (CTBB). Nesses dois momentos de Alencar, analisamos a expressão da forma progressiva (*estar* + *a* + infinitivo *versus* *estar* + gerúndio) e a posição do clítico com o verbo em primeira posição (próclise *versus* ênclise), a fim de verificar se os gêneros em questão influenciam o uso gramatical, ou seja, se há uma relação entre uso da gramática nuclear ou periférica quando o gênero textual é diferente.

A partir dessas considerações, o presente trabalho está organizado da seguinte forma: em §1, discute-se sobre a escolha dos gêneros para a imagem da evolução da língua; em §2, apresentam-se a metodologia empregada e os fenômenos analisados e em §3, são apresentados e discutidos os resultados. Em síntese, este estudo apresenta argumentos que deem suporte à tese de que formas linguísticas usadas pelos falantes na diacronia são influenciadas pelos gêneros textuais, isto é, apresentam variação diagênica⁵.

2 GÊNEROS TEXTUAIS NO TRABALHO DIACRÔNICO

O conjunto de dados com os quais trabalha o historiador da língua envolve a construção de um *corpus*. Faz parte do processo de seleção dos textos considerar aqueles filologicamente

5 Embora existam diferentes termos que fazem referência aos tipos de variação linguística, *viz.* variação diatópica, diastrática, diafásica e diamésica, para se referir à variação regional, social, estilística e de modalidade – oral e escrita –, respectivamente, não há um termo específico para se referir à variação linguística entre gêneros textuais. Neste texto, usamos o termo *variação diagênica* para nos referirmos à variação linguística que tem como principal vetor o gênero textual.

editados, de diferentes períodos e de diferentes gêneros textuais. O gênero textual é um fator importante, uma vez que a escolha por uma ou outra variável por parte do falante/escrevente pode ser a ele condicionada.

Diferentes pesquisadores têm relatado como o fator “gênero textual” influencia os resultados diacrônicos (MATTOS E SILVA, 2008; GALVES, 2018; BIAZOLLI; BERLINCK, 2021; ARAÚJO-ADRIANO; CORÔA, 2022). Por isso, a construção de quaisquer *corpora* necessita de um tratamento mais refinado que controle os gêneros textuais de modo que se possa trabalhar com um *corpus* equilibrado e representativo, evitando o equívoco de se fazer linguística histórica de um único gênero textual. Não é, entretanto, uma tarefa fácil ao historiador da língua reunir textos de diferentes gêneros. Em se tratando da realidade do período colonial do Brasil, o acesso às fontes desse período ainda é muito difícil. Parte dessa massa documental se encontra sob posse do governo português e o que se encontra em solo nacional não goza de uma política de preservação, que ainda é bastante penosa⁶. Além disso, alguns gêneros textuais são muito mais frequentes que outros.

Por exemplo, no *Corpus Tycho Brahe*, contamos com textos portugueses de cartas, atas, textos narrativos, textos dissertativos, gramáticas, gazetas e jornais e textos dramáticos, o que, para um *corpus* diacrônico, é bastante significativo. No que tange à parte brasileira, entretanto, os textos teatrais, as notícias de jornais e as cartas particulares são os textos mais recorrentes e estão concentrados no século XIX, justamente pela política colonialista portuguesa que não permitia a presença de tipografias e universidades (CORÔA, 2022). A multiplicidade de gêneros textuais na constituição de *corpora* diacrônicos permite que possamos melhor entender a variação diagênica, buscando compreender de que modo a gramática nuclear ou a gramática periférica são ativadas a depender do gênero textual utilizado.

Como mostram Araújo-Adriano e Corôa (2022), nem sempre um gênero textual mais próximo da oralidade evidencia as inovações de uma gramática, ou seja, nem sempre é possível capturar nesse ou naquele gênero evidências da diglossia sintática, cuja manifestação ocorre quando uma variante de prestígio e uma variante vernacular competem. As formas em competição podem se diferenciar no registro social, sendo que a variante vernacular corresponde à gramática adquirida na infância e a variante de prestígio é adquirida *via* escolarização, por exemplo.

Também chamada por Kroch (1989; 2001) de *competição de gramáticas*, a diglossia sintática reflete um período de covariação entre opções gramaticalmente incompatíveis na fala e na escrita de um indivíduo e de uma comunidade de fala. Na próxima seção, focamos no indivíduo, especificamente José de Alencar, em dois gêneros textuais distintos, um romance e um texto teatral, a fim de investigarmos o papel desempenhado pelo gênero como resultado da tensão entre uma gramática inovadora e uma gramática conservadora.

3 DOIS GÊNEROS DE JOSÉ DE ALENCAR

3.1 O AUTOR

⁶ Esse cenário mudou um pouco com a iniciativa do Projeto Resgate que “facilitou e expandiu o acesso às fontes para a História do Brasil colonial, ao tratar, microfilmar e digitalizar a documentação concernente às ‘terras de Vera Cruz’ existente no Arquivo Histórico Ultramarino (AHU)” (BERTOLETTI; BELLOTTO; DIAS, 2011, p. 2).

José de Alencar nasceu em 1829, em Messejana, atualmente um bairro de Fortaleza, mas, à época, um município da Província do Ceará. Ainda criança, mudou-se para o Rio de Janeiro, Capital do Império, ambiente retratado em sua produção literária. Alencar é considerado o fundador da literatura brasileira, sendo pioneiro na produção de romances com temática nacional. Em seus romances, sobretudo os da fase indigenista, Alencar revela o processo de formação do povo brasileiro, seus costumes, sua língua, a paisagem e o ambiente da identidade brasileira em formação.

Seguindo a opinião predominante da época de que os teatros no Brasil preferiam peças traduzidas a produções nacionais, escolheu dedicar-se ao teatro como uma forma de promover sua visão de nacionalidade. Entre 1857 e 1862, tornou-se o principal dramaturgo do país. Alencar optou por escrever comédias realistas em vez de dramas românticos, focando nas questões cotidianas e no tempo presente. Seu objetivo era combater as comédias vazias e imorais, buscando unir entretenimento com ensinamentos morais. O autor sempre teve um forte sentimento de nacionalismo, que permeava tanto sua literatura quanto suas peças teatrais. Ele utilizava símbolos e temas como o indianismo, a natureza, o regionalismo e a língua nacional para construir uma identidade brasileira. Além disso, seu projeto teatral visava apontar um caminho para o futuro, reivindicando o progresso e a modernização do Brasil, alinhando-se aos padrões europeus de civilização e costumes urbanos (ver também VASCONCELOS, 2006). Para ele, era essencial estabelecer as origens do nacionalismo e defender a modernização do país.

3.2 O *CORPUS*

Nesta seção, apresentamos o *corpus* sob investigação⁷.

3.2.1 O texto teatral *O demônio familiar* (1857)

Um dia antes da estreia da comédia *O demônio familiar*, escrita em 1857, Alencar define-a como sendo um retrato da vida doméstica brasileira, uma representação dos costumes, um esboço imperfeito das cenas íntimas que ocorrem dentro das casas, e, por fim, uma imagem da família (ver também VASCONCELOS, 2006, p. 174-175). O texto de base utilizado para a análise de dados foi publicado em 1857, pela Typografia Soares & Irmão, portanto a primeira edição do texto. Para levantamento dos dados, as didascálicas foram desconsideradas.

3.2.2 O romance *Iracema* (1865)

Iracema, um dos romances mais reeditados da literatura brasileira, teve sua primeira edição publicada em 1865 e conta com um posfácio – *Carta ao Dr. Jaguaribe* – explicando a

⁷ No catálogo da plataforma *Tycho Brahe*, é possível encontrar outras informações tanto sobre a ficha catalográfica, bem como sobre as edições e o processamento no corpus dos textos utilizados (ver <https://www.tycho.iel.unicamp.br/corpus/catalogo.html>).

gênese e a proposta literária do romance. Este é o texto de base utilizado para a presente análise, na qual foram selecionadas apenas as falas dos personagens.

3.3 BUSCANDO AS GRAMÁTICAS EM JOSÉ DE ALENCAR

Ambos os textos analisados compõem o *Corpus* Histórico do Português Tycho Brahe. O romance *Iracema* ainda não conta com nenhuma anotação, nem morfológica, nem sintática, enquanto o texto teatral *O demônio familiar* conta com a anotação morfológica. Os fenômenos investigados foram (i) a estrutura progressiva e a (ii) colocação de clíticos quando o verbo aparece em primeira posição em uma sentença absoluta.

Uma das diferenças entre o português europeu (PE) e o português brasileiro (PB) emerge na realização da estrutura progressiva. Enquanto na variedade além-mar o progressivo é realizado pelo auxiliar *estar* seguido de uma preposição *a* junto de um domínio infinitivo (*estar + a + infinitivo*), na variedade aquém-mar a mesma estrutura é realizada por *estar* seguido do domínio gerundivo (*estar + gerúndio*), ilustrado abaixo. Nos dados históricos, tal contraste evidenciaria a influência da norma lusitana sobre a língua falada no Brasil.

- (1) a. Olha! Está a chover. (PE, *PB)
b. Olha! Está chovendo. (*PE, PB)

Em uma sentença absoluta, a colocação de clíticos (próclise *versus* ênclise) em primeira posição é uma característica idiossincrática do PB que o diferencia do PE, de tal modo que até foi parar no poema *Pronominais*, de Oswald de Andrade. Nele, o poeta denuncia o contraste entre a ênclise regida pela gramática normativa, herança do PE (*cf.* (2a)), e a próclise, prototípica da nação brasileira (*cf.* (2b)). O contraste analisado neste trabalho visou identificar traços de uma gramática brasileira visto que a próclise é generalizada no PB, enquanto o PE apresenta variação quer em frases finitas quer em frases infinitas (MARTINS, 2018).

- (2) (a) Dê-me um cigarro
Diz a gramática
Do professor e do aluno
E do mulato sabido
Mas o bom negro e o bom branco
Da Nação Brasileira
Dizem todos os dias
Deixa disso camarada
(b) Me dá um cigarro.

A busca dos fenômenos investigados em *O demônio familiar*, texto anotado morfológicamente, foi realizada com o auxílio da ferramenta *CorpusSearch*. Para a estrutura prototípica do PE, *estar + a + infinitivo*, buscamos pela etiqueta ET*, que inclui nos resultados o verbo *estar* em sua forma flexionada, precedendo a preposição *a*, seguida imediatamente por quaisquer verbos no infinitivo (*cf.* (3a)). Para a estrutura *estar + gerúndio*, buscamos pela etiqueta

ET* também precedendo um verbo no gerúndio (cf. (4a)), o que incluíam resultados não só de precedência imediata (cf. (4b)), mas também com algum elemento interveniente (cf. (4c)).

- (3) Formas analíticas
 a. Para *Estar + a + infinitivo*
 (ET* Precedes a)
 AND (a iPrecedes VB|TR|HV|SR)
 b. Está sempre a torcer o bigode!
- (4) a. Para *estar + gerúndio*
 (ET* Precedes VB-G)
 b. Para que procuras esconder uma cousa que teus olhos estão dizendo?
 c. Por sua causa estive quasi não vindo...

Como vai ficar mais claro em §3.2, foi necessário incluímos nas buscas formas sintéticas de presente. Para tanto, buscamos todos os verbos flexionados no presente, representado pela etiqueta VB-P (cf. (5a)). Um exemplo ilustrativo das buscas está apresentado na linha b, a seguir.

- (5) Forma sintética
 a. (VB-P | exists)
 b. Que fazes tu aí? (cf. está fazendo)

Já para o romance *Iracema*, a busca foi feita manualmente, pois, conforme sinalizado anteriormente, o texto ainda não conta com nenhuma anotação. Para tanto, foram seguidos os mesmos contextos rodados na busca automática: foram procurados e contabilizados contextos em que o verbo auxiliar *estar* figurava em uma estrutura *estar + a + infinitivo*, prototípico do PE, e *estar* seguido de gerúndio, podendo ou não haver elementos intervenientes. Para a busca manual das formas sintéticas de presente, foram procurados e contabilizados apenas os contextos em que a forma sintética de presente tinha uma interpretação de simultaneidade em relação à fala⁸, ou seja, contextos em que a forma sintética veiculava a mesma leitura de uma forma analítica estruturada por *estar* (flexionado no presente) + gerúndio, como no caso em que “Não ouves tu, virgem formosa?” é equivalente a “não está ouvindo” (cf. outros casos em (13)).

Para a colocação de clíticos (próclise *versus* ênclise) em primeira posição numa sentença absoluta, as buscas em *O demônio familiar* foram as seguintes: na próclise, buscamos os clíticos que estão em primeira posição, pela etiqueta *isFirst* (cf. (6a)). Nesse caso, desconsideramos as vírgulas em relações de precedência imediata. Na ênclise, buscamos as sentenças absolutas

⁸ Sabe-se que o presente sintético no PB pode também veicular diferentes leituras aspectuais, *viz.* genéricas, habituais, iterativas, prospectivas *etc* (ver também MIOTO, 1985; FATORI, 2010; BARBOSA; CRUZ, 2013; ARAÚJO-ADRIANO, 2023). Porém, neste trabalho, foi considerada apenas a leitura em que a forma sintética veicula uma noção temporal de presente, isto é, uma noção de simultaneidade presente em relação à fala.

cujos verbos aparecem em primeira posição (cf. (7a)). Alguns exemplos estão apresentados nas linhas b, a seguir.

- (6) Próclise
a. CL|SE|CL+*|SE+* isFirst
b. Te arrependeste a tempo.
- (7) Ênclise
a. VB*|ET*|TR*|SR*|HV*+CL isFirst
b. Esquecia-me que não gostas que adivinhem os teus segredos.

Já para o romance *Iracema*, a busca foi feita manualmente, de acordo com os mesmos contextos rodados na busca automática, listados em (6) e em (7).

4 AS DUAS GRAMÁTICAS DE JOSÉ DE ALENCAR

Nesta seção, apresentamos os resultados obtidos a partir da investigação do posicionamento do clítico e da estrutura progressiva tanto no romance quanto na peça teatral de José de Alencar. Em seguida, fornecemos uma análise dos principais resultados, com o objetivo de responder às questões delineadas na introdução.

4.1 AS DUAS GRAMÁTICAS DE JOSÉ DE ALENCAR: O CLÍTICO QUANDO O VERBO ESTÁ NA PRIMEIRA POSIÇÃO NA SENTENÇA

Pela análise do posicionamento do clítico nos dois gêneros, foi possível observar que o uso da ênclise é categórico em ambos os textos: 100% no romance (cf. (9)) e 99% no texto teatral (cf. (8)). Ainda que timidamente, a forma típica brasileira, isto é, próclise com o verbo na primeira posição na sentença, apenas é atestada no texto teatral.

- (8) Alencar no texto teatral (1857)
Ênclise (99%)
a. Esquecia-me que não gostas que adivinhem os teus segredos.
b. Mostre-se arrufado que é para ela responder.
c. Abro-o, e é uma carta de namoro!
- Próclise (1%)
d. Te arrependeste a tempo.
- (9) Alencar no romance (1865)
Ênclise (100%)
a. Mandou-me a ti para ouvir
b. Arreda-te do ennemigo, virgem dos Tabajaras

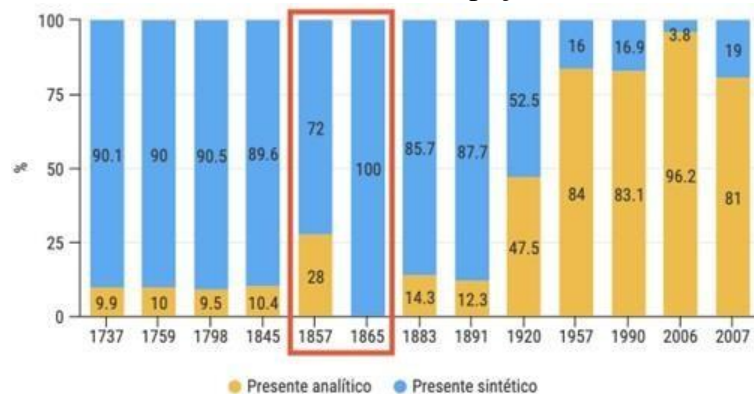
c. Escrevi-o para ser lido

A seguir, apresentamos o Gráfico 1 comparando a ocorrência do clítico quando o verbo está na primeira posição na sentença em diversos autores brasileiros a partir da data de publicação dos textos. No gráfico, tanto a análise dos textos de Alencar (1857; 1865) quanto dos demais autores brasileiros (demais períodos) reflete uma tendência no português do Brasil de manutenção da ênclise em primeira posição⁹.

Adotando Pagotto (1998), seguimos a hipótese de que a influência da norma portuguesa parece mascarar a ocorrência da forma brasileira. Para reforçar a hipótese levantada, ressalta-se que Alencar, enquanto integrante da elite intelectual do país naquele período, espelha a variedade de prestígio do português brasileiro, bem como os demais autores analisados no Gráfico 1. Adicionalmente, Carneiro (2005) aponta que a ocorrência de próclise em primeira posição absoluta, sobretudo no início do século XIX, estaria associada a uma vertente menos prestigiada do português brasileiro. Em sua tese, a ocorrência da próclise quando o verbo está na primeira posição na sentença é maior entre os escreventes semicultos do que entre escreventes cultos.

Ainda, o Gráfico 1 evidencia uma mudança fracassada (PAGOTTO, 1992; CARNEIRO, 2005), visto que a partir da segunda metade do século XX a próclise se impõe como tendência natural da língua.

Gráfico 1 – Ênclise ~ Próclise em peças de teatro e no Romance



Fonte: elaborado pelos autores.

Seguindo a hipótese levantada neste trabalho, esperaríamos que houvesse diferença significativa entre os gêneros textuais, com o romance apresentando um número muito maior de formas conservadoras (ênclise) e o texto teatral com mais formas inovadoras (próclise). O

⁹ Galves, Britto e Paixão de Sousa (2005) defendem que, na gramática do português clássico, a ênclise é licenciada apenas nos contextos em que o verbo está na primeira posição na oração. A restrição seria a seguinte: um clítico não pode ser o primeiro elemento do primeiro XP da oração. Essa restrição se correlaciona com a Lei Tobler-Mussafia, que impede que um item não acentuado apareça no início de uma oração. Aplicada à sentença, a lei mencionada prevê a ocorrência de ênclise exclusivamente em contextos em que o verbo finito é o primeiro elemento da estrutura prosódica da oração e prevê também que a próclise ocorra nos demais contextos.

que vemos é a prevalência da ênclise sobre a próclise, atestada não apenas em Alencar, mas em quaisquer escritos ao longo do século XIX. É apenas na segunda metade do século XX que a ênclise dá lugar à próclise. O que, então, explicaria esse padrão?

Na segunda metade do século XIX, há uma grande onda purista no Brasil que toma a gramática portuguesa como norma para o português do Brasil (GUIMARÃES, 1994). Desse modo, o século XIX termina “com a norma culta mais europeia do que nunca” (PAGOTTO, 1998, p. 56), como pode ser notado, por exemplo, nas gramáticas normativas da época, que adotam a norma lusitana para uma série de fatos linguísticos, incluindo a colocação dos pronomes clíticos.

Pagotto (1998) enfatiza que, ao longo do século XIX, a forte pressão normativa fez com que fala e escrita seguissem caminhos opostos e que só começaram a estar mais *pari passu* na segunda metade do século XX, como atestado no Gráfico 1. A oposição entre a fala e a escrita, fruto da forte pressão da norma lusitana nos textos brasileiros (PAGOTTO, 1998; CARNEIRO, 2005; LOBO, 2001), é atestada em grandes polêmicas nos jornais da época, tais como as polêmicas entre Ernesto Carneiro Ribeiro *versus* Rui Barbosa e Pinheiro Chagas *versus* José de Alencar.

Apesar da onda purista, Alencar era um dos escritores que levantava a bandeira em favor de uma língua brasileira e por isso foi alvo de diversas críticas. Uma das maiores polêmicas em que se envolveu em favor de sua posição nacionalista foi contra as críticas do português Pinheiro Chagas que se dirigiram ao romance *Iracema*, como visto na citação a seguir:

o defeito que eu vejo nessa lenda [Iracema], o defeito que vejo em todos os livros brasileiros, contra o qual não cessarei de bradar intrepidamente, é a falta de correção na linguagem portuguesa, ou antes a mania de tornar o brasileiro uma língua diferente do velho português, por meio de neologismos arrojados e injustificáveis, e de insubordinações gramaticais, que, (tenham cautela!) chegarão a ser risíveis se quiserem tomar as proporções de uma insurreição em regra contra a tirania de Lobato (CHAGAS, 1867, p. 212, grifo nosso).

As críticas de Pinheiro Chagas direcionam-se à presença de neologismos e de certas construções gramaticais que ele nomeia como insubordinações. De acordo com Guimarães (1994), também são feitas críticas ao uso do artigo, omissão do *se* reflexivo de certos verbos, troca do *se* pelo *si*, além da colocação de pronomes pessoais.

Apesar das críticas, Alencar seguia muito estritamente a norma europeia no que tange à colocação do clítico quando o verbo está na primeira posição na sentença. Como vimos, em *Iracema*, o uso da ênclise é categórico; a forma brasileira apenas aparece no texto teatral. Isso evidencia o poder que a pressão normativa exercia sobre os escritores do período, mesmo entre aqueles que advogavam por uma gramática brasileira, como Alencar. Também fica patente que a pressão normativa atuou entre os gêneros textuais mais próximos à oralidade, como vemos no Gráfico 1, cujos textos, com exceção do romance *Iracema* (1865), são todos peças de teatro.

Seguindo a discussão levantada por Kroch (1989; 2001), os padrões atestados em Alencar evidenciam o processo de competição de gramáticas. A competição de gramáticas é fruto da tensão entre a aquisição natural da língua e sua aprendizagem por meio do ensino

formal e revela que “as formas em competição podem se diferenciar no registro social, com uma variante vernacular lentamente guiando uma variante escrita fora de uso” (KROCH, 2004, p. 6).

Essa tensão é atestada não só nas polêmicas entre Alencar e Pinheiro Chagas mas também no próprio texto teatral, gênero mais próximo da oralidade. Por um lado, vemos os padrões de uma gramática enclítica, marcada normativamente na escrita; por outro, vemos os padrões de uma gramática proclítica, inovadora, similar à gramática do PB, usada na fala (ver discussão sobre as polêmicas envolvendo Alencar e Pinheiro Chagas, mas também PAGOTTO, 1998; CARNEIRO, 2005). Assim teríamos:

- (10) a. Gramática Nuclear: padrão proclítico
- b. Gramática Periférica: padrão enclítico

4.2 AS DUAS GRAMÁTICAS DE JOSÉ DE ALENCAR: A ESTRUTURA PROGRESSIVA

Inicialmente, a partir de diferentes gêneros textuais, o presente trabalho buscou encontrar evidências da tensão entre uma gramática conservadora lusitana e uma gramática inovadora puramente brasileira. Considerando isso, a proposta era investigar como o gênero textual influenciava fenômenos prototípicos de uma variedade ou outra. Esse foi o caso do posicionamento clítico, *supra* apresentado, em que o clítico em primeira posição é característico da variedade brasileira diferentemente da variedade lusitana, que realiza o clítico nesse contexto em posição pós-verbal. Quanto aos resultados da estrutura progressiva, fomos surpreendidos com um fato que pode revelar a influência da variação diagênica na tensão entre uma gramática inovadora e uma gramática conservadora.

Quando comparados os dois gêneros, os dados referentes à estrutura progressiva variaram drasticamente no mesmo autor; enquanto nas peças teatrais (1857) obtivemos 3% das ocorrências com a forma progressiva formada por *estar + a + infinitivo* e 97% com *estar + gerúndio*, exemplificadas em (11), no romance (1865) não obtivemos nenhuma forma progressiva.

- (11) Alencar no texto teatral (1857)
 - a. Forma progressiva com *estar + a + infinitivo*
a'. Está sempre a torcer o bigode!
 - b. Forma progressiva com *estar + gerúndio*
b'. Mana Carlotinha, Henriqueta está lhe chamando para dizer-lhe adeus.
b". Tu estais tirando os charutos de mano!

A surpresa fundamenta-se em duas expectativas gerais, quais sejam: (i) a de que no romance haveria um número muito maior de formas conservadoras – as formas progressivas lusitanas – ou (ii) a de que o romance apresentaria mais formas inovadoras, por ser escrito após a peça teatral, isto é, mais próxima da gramática contemporânea. Apesar de não haver estruturas analíticas de Presente com gerúndio no romance, José de Alencar fazia uso de

formas gerundivas, seja em sentenças adjuntas (12a)-(12b), seja em outras estruturas perifrásticas (12c).

- (12) a. Teu hóspede espera, filha de Araken: mas se o sol tornando, não trouxer o irmão de Iracema, ele levará o guerreiro branco á taba dos Pytiguaras.
b. Mais amigo e companheiro será de Poty, servindo a seu irmão que a ele.
c. Vêem teus olhos lá o formoso jacarandá, que vai subindo ás nuvens;

Dada a quebra de expectativa, a pergunta que imediatamente emergiu foi a seguinte: “se não são usadas formas progressivas nem conservadoras tampouco inovadoras, como a expressão de Tempo presente é codificada no romance?”. Para responder a essa pergunta, baseamo-nos na discussão em Araújo-Adriano (2023). Investigando a expressão do Tempo presente no PB, o autor mostra que, na contemporaneidade, falantes usam categoricamente a forma progressiva para veicular um evento simultâneo à fala, isto é, Tempo presente. Assim, ao observar água caindo no céu, é provável que um falante de PB diga “Olha, **está chovendo**”, mas não a forma sintética equivalente “Olha, **chove**”. Na diacronia da língua, entretanto, Araújo-Adriano (2023) nota que falantes de PB usavam majoritariamente a forma sintética para veicular tal noção temporal, ainda que a forma analítica, inovadora, também fosse usada (mais timidamente).

Voltando a José de Alencar e à nova pergunta de pesquisa, no romance escrito em 1865, dada a ausência de formas analíticas de presente, *supra* atestada, 100% dos casos para veicular Tempo presente foram atestados com formas sintéticas (13). Isso significa que Alencar recorria a uma gramática completamente conservadora, não abrindo espaço para sequer uma forma inovadora no romance. Esse padrão, entretanto, não era categórico em outros gêneros.

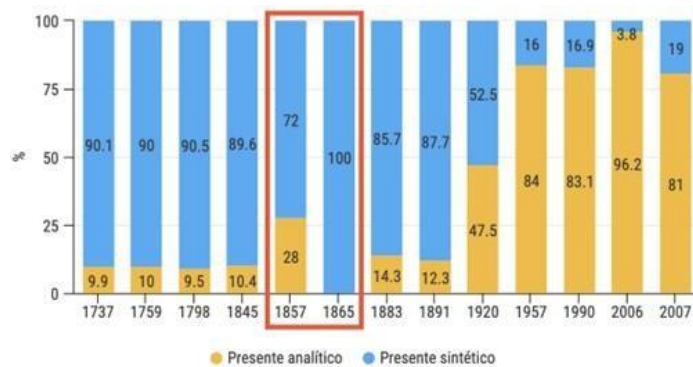
- (13) Alencar no romance (1865)
a. Não ouves tu, virgem formosa? exclamou ele apontando para o antro fremente. (*cf.* estás ouvindo)
b. O cristão caminhou para a cabana de Jacaúna. O grande chefe alegrou-se vendo chegar seu hóspede; mas a alegria fugiu logo de sua fronte guerreira. Martim dissera: O guerreiro branco parte de tua cabana, grande chefe. (*cf.* está partindo)
c. Guerreiro Poty, teu irmão branco te chama pela boca de Iracema. (*cf.* está te chamando)

Contrariamente ao encontrado no romance, na peça teatral escrita oito anos antes do romance, José de Alencar usava uma gramática com 28% das ocorrências com a forma inovadora (*cf.* (11)), em contraste com 72% com a conservadora (14), isto é, a forma sintética. Apesar de a forma inovadora ser pouco recorrida, esses resultados sugerem que ela se faz presente apenas na peça, sendo completamente ignorada no romance.

- (14) Peça de teatro (1857)
a. Em que pensas, Eduardo? (*cf.* estás pensando)
b. Que fazes tu aí? (*cf.* está fazendo)
c. Henriqueta te chama, Carlotinha. (*cf.* está te chamando)

Ainda sobre a diacronia das expressões de Tempo presente, Araújo-Adriano (2023) atesta que a preferência em peças teatrais pela forma analítica aumenta ao longo do tempo, passando a ser a forma padrão a partir do fim do século XIX. Posto ao lado do gênero teatro, escrito em 1857, o gênero romance, mesmo escrito mais tardiamente, em 1865, interrompe essa tendência atestada de modo geral, como se observa no Gráfico 2. Essa quebra na tendência natural da língua justifica as expectativas gerais e a surpresa nos resultados supra levantados.

Gráfico 2 – Presente analítico e sintético em peças de teatro e no Romance



Fonte: elaborado pelos autores

Considerando que a forma analítica de presente foi completamente negligenciada no romance, um texto menos oralizado, por hipótese essa forma pode ser concebida como sendo a forma marcada, em contraste com a forma sintética não marcada, vernacular. Em termos de competição de gramática (KROCH, 1989), o mesmo autor – José de Alencar – teria duas, tais quais formuladas em (15) abaixo. O acionamento de uma em detrimento da outra estaria, discutivelmente, condicionado à variação diagênica.

- (15) a. Gramática Nuclear: forma sintética de presente (gênero mais formal)
b. Gramática Periférica: forma analítica de presente (gênero menos formal)

5 CONCLUSÃO

Este trabalho discutiu a influência dos gêneros textuais na ativação da gramática nuclear ou periférica de um mesmo autor e, em última instância, de um mesmo falante. A partir da comparação entre dois textos de José de Alencar, o romance *Iracema* e a peça teatral *O demônio familiar*, a hipótese de trabalho era a de que o gênero romance condicionasse a gramática periférica, emergindo formas conservadoras, enquanto o gênero teatro condicionaria a gramática nuclear, mantendo formas inovadoras.

Os resultados preliminares confirmam em parte a hipótese levantada. Quando o verbo está em primeira posição na sentença, o uso da ênclise é categórico em ambos os textos: 100% no romance e 99% no texto teatral. Tanto no romance quanto na peça teatral, vê-se

representada a gramática periférica do autor (cf. (16)), embora a próclise fosse a colocação clítica da gramática nuclear. Quanto à expressão de Tempo presente, a forma inovadora (*estar* + gerúndio), manifestada pela gramática usada no gênero teatral em 28% das variantes, não é sequer testemunhada no gênero romance, sendo extensivamente veiculada pela forma sintética de presente, a forma conservadora.

- (16) a. Gramática Nuclear: padrão proclítico, forma sintética de presente
b. Gramática Periférica: padrão enclítico, forma analítica de presente

Concluimos com este trabalho que o mesmo autor faz uso de gramáticas distintas, a depender do gênero textual, seguindo a proposta de Kroch (2001, p. 6) de competição de gramáticas (diglossia sintática) que ocorre quando “as formas em competição podem se diferenciar no registro social, com uma variante vernacular lentamente guiando uma variante escrita fora de uso”. Esta investigação não se esgota aqui. Pretendemos ampliá-la, contrastando gêneros textuais distintos em outros autores, além de tentar recuar temporalmente, a fim de evidenciar a variação diagênica em outros estágios da língua.

REFERÊNCIAS

- ALENCAR, J. de. *Iracema: Lenda do Ceará*. Rio de Janeiro, Typ. De Vianna & Filhos, Rua D'Ajuda n° 79, 1865.
- ALENCAR, J. de. *O demonio familiar: comedia em quatro actos*. Rio de Janeiro: Typografia Soares & Irmão, 1857.
- ARAÚJO-ADRIANO, P. Â. Conhecimento linguístico do letrado acerca das expressões de futuridade: Forma perifrástica (gramática nuclear/L1) versus forma sintética (gramática periférica/L2). *Letras de Hoje*, [s. l.], v. 55, n. 3, p. 300–315, 2020.
- ARAÚJO-ADRIANO, P. Â. *The Present Tense Analyticisation Process in Brazilian Portuguese: A Diachronic Approach*. 2023. 293f. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, Campinas, SP, 2023.
- ARAÚJO-ADRIANO, P. Â.; CORÔA, W. Reconstruindo a história do português do Brasil pelo corpus Tycho Brahe Brasil: novos dados, novos olhares. *Revista Linguística*, [s. l.], v. 17, n. 3, p. 202–227, 2022.
- BARBOSA, J. B.; CRUZ, R. C. V. Os valores semânticos do presente do indicativo no português brasileiro: um estudo em blogs. *Entretextos*, Londrina, v. 13, n. 1, p. 53–79, 2013.
- BIAZOLLI, C. C.; BERLINCK, R. Por que investigar processos de variação e mudança linguísticas por meio de gêneros textuais-discursivos?. In: BIAZOLLI, C. C.; BERLINCK, R. de A. (org.). *Gêneros textuais-discursivos no estudo de processos de variação e mudança*. Campinas: Pontes, 2021. p. 13–38.
- CARNEIRO, Z. O. N. *Cartas brasileiras (1809-1904): um estudo linguístico-filológico*. 2246 f. Tese (Doutorado em Linguística) — Universidade Estadual de Campinas (UNICAMP). Instituto de Estudos da Linguagem, Campinas, 2005.

- CHAGAS, P. Literatura brasileira – José d’Alencar. In: *NOVOS ENSAIOS CRÍTICOS*. Porto: Casa da viúva Moré, 1867. p. 212–213.
- CHOMSKY, N. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger, 1986.
- CORÔA, W. *Rastreado as origens do Português Brasileiro: a dinâmica da mudança na escrita de “Homens Bons” na Bahia Colonial*. 2022. 406 f. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, Campinas, SP, 2022.
- CYRINO, S.; TORRES MORAIS, M. A. (org.). *Mudança sintática do português brasileiro: perspectiva gerativista*. São Paulo: Contexto, 2018. v. 6 História do Português Brasileiro (Coordenador geral: Ataliba de Castilho).
- FATORI, M. J. *Um estudo semântico-discursivo sobre o emprego do presente do indicativo no português do Brasil*. 2010. Tese (Doutorado em linguística e língua portuguesa) - UNESP - Universidade Estadual Paulista, Araraquara, SP, 2010.
- GALVES, C. Posfácio: o retrato da emergência de uma nova gramática. In: CYRINO, S.; TORRES MORAIS, M. A. (org.). *Mudança sintática do português brasileiro: perspectiva gerativista*. São Paulo: Contexto, 2018. v. 6 História do Português Brasileiro (Coordenador Geral: Ataliba de Castilho), p. 441–456.
- GALVES, C.; BRITO, 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*, [s. l.], v. 4, p. 39–57, 2005.
- GALVES, Charlotte; ANDRADE, Aroldo Leal de; FARIA, Pablo (2017, Dezembro). *Tycho Brahe Parsed Corpus of Historical Portuguese*. URL: texts/psd.zip.
- KABATEK, J. Tradições discursivas e mudança lingüística. In: ALKMIN, T. M. et al. (org.). *Para a história do português brasileiro*. Salvador: EDUFBA, 2006. v. 6 Novos dados, novas análises (Tomo II), p. 505–527.
- KATO, M. A. A gramática nuclear e a língua-i do brasileiro. In: MARTINS, M. A. (org.). *Gramática e ensino*. Natal: Editora EDUFRN, 2013. (Coleção Ciências da Linguagem Aplicadas ao Ensino, v. 1). p. 145–159.
- KATO, M. A.; CYRINO, S. L.; CORRÊA, V. Brazilian Portuguese and the recovery of lost clitics through schooling. In: PIRES, A.; ROTHMAN, J. (org.). *Minimalist Inquiries into Child and Adult Language Acquisition*. New York: Mouton de Gruyter, 2009. (Studies on Language Acquisition). v. 35, p. 245–272.
- KATO, M. Aquisição e aprendizagem da língua materna de um saber inconsciente para um saber metalingüístico. In: *Investigações à linguagem: ensaios em homenagem a Leonor Scliar-Cabral*. Florianópolis: Mulher, 1999. p. 201–205.
- KROCH, A. S. Reflexes of grammar in patterns of language change. *Language Variation and Change*, [s. l.], v. 1, n. 3, p. 199–244, 1989.
- KROCH, A. S. Syntactic Change. In: BALTIN, M.; COLLINS, C. (ed.). *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell Publishers, 2001. p. 699–729.
- LIGHTFOOT, D. *Principles of diachronic syntax*. Cambridge, UK: Cambridge University Press, 1979.
- LOBO, Tânia. *Para uma sociolinguística histórica do português no Brasil: edição filológica e análise linguística de cartas particulares do Recôncavo da Bahia, século XIX*. 2001. 808 f. Tese (Doutorado em Filologia e Língua Portuguesa) - Universidade de São Paulo, São Paulo, 2001.

- MARTINS, Marco Antonio Rocha. A sintaxe dos pronomes pessoais clíticos na história do português brasileiro. In: Cyrino, Sonia & Maria A. Torres Morais. (Org.). *Mudança sintática do português brasileiro: perspectiva gerativista*. 1ed. São Paulo: Contexto, 2018, v. IV, p. 150-209.
- MATTOS E SILVA, R. V. *Caminhos da linguística histórica: ouvir o inaudível*. São Paulo: Parábola, 2008.
- MIOTO, C. Considerações sobre o presente do indicativo. In: SEMINÁRIO DO GRUPO DE ESTUDOS LINGÜÍSTICOS DE SÃO PAULO, 1985, Faculdade do Sagrado Coração. *Grupo de Estudos Linguísticos de São Paulo*. Faculdade do Sagrado Coração: [s. n.], 1985. p. 16–21.
- PAGOTTO, E. G. *A posição dos clíticos em português: um estudo diacrônico*. 1992. 157 f. Dissertação (Mestrado em Linguística) - Instituto de Estudos da Linguagem, Unicamp, Campinas, 1992.
- PAGOTTO, E. G. Norma e Condescendência; Ciência e Pureza. *Línguas Instrumentos Linguísticos*, [s. l.], v. 2, p. 49–68, 1998.
- PINTO, C. F. C. ; ANDRADE, A. L.. Desmistificando a gramática gerativa como uma teoria associal e a-histórica da mudança linguística. *MACABÉA- REVISTA ELETRONICA DO NETLLI*, v. 8, p. 36-66, 2019.
- ROMAINE, Suzanne. *Socio-historical linguistics: its status and methodology*. Cambridge: Cambridge University Press, 1982.
- TARALLO, Fernando. *A pesquisa Sociolinguística*. 8 ed. São Paulo: Ática, 2007.
- VASCONCELOS, P. H. O teatro de José de Alencar e a imagem do negro no Brasil do século XIX. *GUANICUNS: Rev. Faculdade de Educação e Ciências Humanas de Anicuns FECHA/FEA*, Goiás, n. 3/4, p. 167–181, 2006.
- WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory of language change. In: LEHMANN, W.; MALKIEL, Y. (ed.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p. 95–188.