

A variação diacrônica entre *que* e *o que* interrogativos em peças teatrais brasileiras e portuguesas

DOI: <http://dx.doi.org/10.21165/el.v54i3.3850>

Vitor Morelli Silva¹

Resumo

Este artigo discute a variação diacrônica entre “que” e “o que” interrogativos no português brasileiro em comparação com a que ocorre no português europeu. Sob uma ótica gerativista, a sintaxe diacrônica (Roberts, 2021), trabalha-se com dados de interrogativas diretas e indiretas obtidas de um conjunto de 17 obras de teatro do Corpus Histórico do Português Tycho Brahe, de autores brasileiros nascidos entre os séculos XVIII e XX e autores portugueses nascidos entre os séculos XVIII e XIX. Trabalha-se a partir da hipótese de que, na segunda metade do século XIX no Brasil, a tendência de aproximação da gramática do português europeu, observada em Pagotto (1992, 2013, 2018) e Carneiro (2005), também pode ser identificada nesta variação. Argumenta-se a favor dessa hipótese com base nos resultados quantitativos confrontados com as prescrições de gramáticos do século XIX.

Palavras-chave: o que; que; variação; orações interrogativas.

¹ Universidade Estadual de Campinas (Unicamp), Campinas, São Paulo, Brasil; vitimorellis06@gmail.com; <https://orcid.org/0000-0002-9596-240X>

The diachronic variation between interrogative *que* and *o que* in Brazilian and European Portuguese theatre

Abstract

This article discusses the diachronic variation between interrogative “que” and “o que” in Brazilian Portuguese in comparison with that occurring in European Portuguese. From a generativist perspective, diachronic syntax (Roberts, 2021), this study works with data from direct and indirect questions obtained from a set of 17 dramatic works from the Tycho Brahe Parsed Corpus of Historical Portuguese, by Brazilian authors born between the 18th and 20th centuries and Portuguese authors born between the 18th and 19th centuries. The work is based on the hypothesis that, in the second half of the 19th century in Brazil, the tendency towards approximation to European Portuguese grammar, observed in Pagotto (1992, 2013, 2018) and Carneiro (2005), can also be identified in this variation. This hypothesis is argued on the basis of quantitative results compared with the prescriptions of 19th-century grammarians.

Keywords: *o que*; *que*; variation; interrogative clauses.

Introdução

Emergiu na língua em Portugal, entre os séculos XVII e XVIII, um novo pronome interrogativo, *o que*, que entraria em variação com *que* (Silva, 2023; Galves; Andrade, Silva, 2026, no prelo). Galves, Andrade e Silva (2026, no prelo) argumentam que a entrada de *o que* no quadro de pronomes interrogativos do português seria o resultado do processo de gramaticalização por que passou *o que* entre os séculos XVII e XVIII. Isso se encaixa, pode-se dizer, num conjunto de mudanças sintáticas e prosódicas por que passou o português nesse mesmo período e que culminou no que hoje conhecemos como o português europeu (cf. Galves; Paixão de Sousa, 2017). São exemplos dessa variação:

1. *O que* fará o meu bem? (Rodrigues da Costa, nascido no séc. XVIII)
2. E você, *que* diz? (Rodrigues da Costa, nascido no séc. XVIII)

A mesma variação é atestada em orações interrogativas indiretas:

3. porém, perguntando-me Vossa Senhoria *o que* penso dele, digo-lhe que o caso para mim tem pouco de extraordinário. (Cavaleiro de Oliveira, nascido no séc. XVIII)
4. Senhor Careiro, eu mandei-o chamar porque quero ajustar contas e saber *que* lhe devemos de cinco meses *que* aqui estamos. (Rodrigues da Costa, nascido no séc. XVIII)

Mateus *et al.* (2006, p. 464, grifo dos autores), em sua gramática de base gerativista do português europeu (PE), afirma que a forma *que* é cada vez menos utilizada, “sendo *o que* a forma mais produtiva”. É notório que uma gramática contemporânea do PE já reconheça *que* interrogativo como uma forma antiga. Nos *corpora* Corpus Histórico do Português Tycho Brahe² e P. S. Post Scriptum,³ conforme Galves, Andrade e Silva (2026, no prelo) e Silva (2023),⁴ observa-se em textos de autores portugueses (de diversos gêneros no primeiro e de cartas no segundo) que as ocorrências de *o que*, em orações interrogativas diretas, aparecem no século XVIII e aumentam sutilmente dos séculos XVIII a XIX (Tabela 1). Em orações interrogativas indiretas, *o que* já aparece como preferido desde o século XVI (Tabela 2).

Tabela 1. Ocorrências de *que* e *o que* e porcentagens de *o que* por séculos em interrogativas diretas em textos de autores portugueses (Corpus Tycho Brahe e Corpus Post Scriptum)

Século	<i>Que</i>	<i>O que</i>	Frequência relativa de <i>o que</i>
XVI	8	0	0%
XVII	67	0	0%
XVIII	172	5	2,82%
XIX	128	12	8,57%

Fontes: Galves, Andrade e Silva (2026, no prelo) e Silva (2023, p. 27)

Tabela 2. Ocorrências de *que* e *o que* e porcentagens de *o que* por séculos em interrogativas indiretas (Corpus Tycho Brahe e Corpus Post Scriptum)

Século	<i>Que</i>	<i>O que</i>	Frequência relativa de <i>o que</i>
XVI	7	18	72%
XVII	6	32	84,21%
XVIII	11	12	52,17%
XIX	2	55	96,49%

Fonte: Silva (2023, p. 28-29)

Fala-se aqui em “interrogativas indiretas” porque as orações em questão foram anotadas conforme as regras de anotação para interrogativas indiretas.⁵ No entanto, Galves,

2 GALVES, C.; ANDRADE, A. L. de; FARIA, P. *Tycho Brahe Parsed Corpus of Historical Portuguese*. Dez. 2017. Disponível em: texts/psd.zip. Acesso em 4 jul. 2024.

3 CLUL (ed.). *P.S. Post Scriptum*. Arquivo Digital de Escrita Quotidiana em Portugal e Espanha na Época Moderna. 2014. Disponível em: <http://ps.clul.ul.pt>. Acesso em: 4 jul. 2024.

4 Conforme apontado por Galves, Andrade e Silva (2026, no prelo), em Silva (2023), há um dado com *o que* no século XVII cuja interpretação é difícil. Numa análise mais minuciosa do contexto da oração, concluiu-se que, na realidade, se trata de uma oração relativa livre.

5 Cf. MAGRO, C.; GALVES, C. *Portuguese Syntactic Annotation Manual*. 2019. Disponível em: <http://alfclul.clul.pt/portuguesesyntacticannotation/home.html>. Acesso em: 4 jul. 2024.

Andrade e Silva (2026, no prelo) argumentam que, antes do século XVII, apenas havia “relativas livres” com *o que*, em que *o* tinha valor dêitico e selecionava uma oração relativa como complemento, podendo ser interpretadas tanto como interrogativas indiretas como relativas livres, semelhante ao que ocorre com estruturas com *ce que* em francês.

No português brasileiro (PB), também se observa variação (em textos recentes, aliás) entre *que* e *o que* em orações interrogativas diretas, como em (5) e (6):

5. a. *O que* vai ser dele? (Paulo Gustavo A. M. de Barros, nascido no séc. XX)
b. *Que* foi, mãe? (Paulo Gustavo A. M. de Barros, nascido no séc. XX)
6. a. *O que* elas vão pensar de mim? (Paulo Sacaldassy, nascido no séc. XX)
b. *Que* você fez? (Paulo Sacaldassy, nascido no séc. XX)

Neste artigo,⁶ discute-se a variação entre *que* e *o que* interrogativos no PB em comparação com o PE, dando-se especial atenção à relação dos autores brasileiros com a norma. Os trabalhos sobre a colocação de clíticos no PB de Pagotto (1992, 2013, 2018) e Carneiro (2005) observam, na segunda metade do século XIX, uma tendência de adoção do padrão do PE. Partindo da hipótese de que essa tendência também pode ser observada em outras estruturas gramaticais em variação, este trabalho investiga se esse fenômeno também pode ocorrer para a variação entre *que* e *o que* em orações interrogativas. A consulta às gramáticas permite investigar se essa influência do PE também se dá por meio de pressão normativa. Trabalha-se aqui com um conjunto de 17 obras teatrais do Corpus Tycho Brahe: 6 portuguesas (Tabela 3) e 11 brasileiras (Tabela 4).⁷ A escolha do teatro se justifica por este ser um gênero em que há orações interrogativas em quantidade mais elevada, como já atestado por Lopes-Rossi (2018). Considerando o critério do ano de nascimento, este conjunto contempla os séculos XVIII e XIX para Portugal⁸ e XVIII a XX para o Brasil. Porém, considerando o critério da data de publicação, os textos portugueses contemplam os séculos XIX e XX⁹ e os textos brasileiros contemplam os séculos XVIII a XXI. Os dados foram coletados por meio de buscas morfológicas e sintáticas¹⁰ com

6 Este artigo deriva parcialmente da dissertação de mestrado do autor, com defesa agendada para 31 de março de 2026, no Instituto de Estudos da Linguagem (Unicamp).

7 As informações sobre as obras apresentadas nas Tabelas 3 e 4 foram retiradas das fichas catalográficas do Corpus Tycho Brahe.

8 Esta pesquisa fez uso exclusivamente das obras que já fazem parte do Corpus Tycho Brahe, por esta razão o século XX não foi incluído em Portugal pelo critério do ano de nascimento do autor.

9 A obra teatral *6 Entremeses de Cordel*, do autor português José Daniel Rodrigues da Costa, nascido em 1757, é uma coletânea publicada em 1973 (cf. Tabela 3), que, por sua vez, se baseia no terceiro volume da publicação *Teatro Cômico de Pequenas Peças*, do mesmo autor, cuja edição referenciada é de 1879. É informado no prefácio que houve atualizações na ortografia e na pontuação (cf. Costa, 1973, p. 27). Contudo, não se espera que essas atualizações afetem as estruturas aqui visadas.

10 As buscas sintáticas foram rodadas apenas nos textos do século XVIII, pois são os que já têm anotação sintática completa.

Corpus Search¹¹ e na Plataforma Tycho Brahe.¹² Além disso, foram feitas consultas aos próprios textos para tentar evitar a perda de dados.¹³ Foram selecionadas para análise apenas ocorrências em contextos oracionais.¹⁴

Tabela 3. O *corpus* português

Autor	Nascimento	Publicação (original/fonte)	Título da obra
José Daniel R. da Costa	1757	1973 (fonte)	<i>6 Entremezes de Cordel</i>
Joaquim da C. Cascais	1815	1899	<i>A inauguração da estátua equestre</i>
Camilo C. Branco	1825	1871	<i>A Morgadinha de Val-d'Amores</i>
Manuel P. Chagas	1842	1869	<i>A Morgadinha de Valflor</i>
Marcelino Mesquita	1856	1897	<i>O Regente</i>
André Brun	1881	1916	<i>Teatro, André Brun</i>

Fonte: Elaboração própria

Tabela 4. O *corpus* brasileiro

Autor	Nascimento	Publicação (original/fonte)	Título da obra
Antonio José da Silva	1705	1759 (fonte)	<i>Theatro comico portuguez I</i>
Luís Carlos M. Pena	1815	1838/1871	<i>O juiz de paz na roça</i>
Luís Carlos M. Pena	1815	1845/1853	<i>O noviço: comedia em 3 actos</i>
José de Alencar	1829	1858	<i>O demônio familiar: comedia em quatro actos</i>

11 RANDALL, B. *CORPUS SEARCH 2: a tool for linguistic research*. 2004 Disponível em: <http://corpussearch.sourceforge.net/CS.html>. Acesso em: 11 set. 2025.

12 VERONESI, L. H. L.; GALVES, C. *The Tycho Brahe Platform*. 2024. Disponível em: <https://www.tycho.iel.unicamp.br/home>. Acesso em: 11 set. 2025.

13 As buscas morfológicas foram feitas na versão *on-line* de Corpus Search no *site* do Corpus Tycho Brahe e, por serem mais simples, seus resultados nem sempre contêm apenas dados relevantes, de modo que precisaram ser filtrados e analisados manualmente. A Plataforma Tycho Brahe foi utilizada apenas para encontrar as interrogativas indiretas, pois a busca por essas estruturas foi mais bem-sucedida na plataforma.

14 Casos em que a palavra-QU precede um advérbio, *e.g.*, “Que/O que mais...?”, somente foram quantificados quando havia um verbo na frase. Casos de “Que isso?” foram considerados equivalentes a “Que é isso?” e, portanto, também foram quantificados.

Joaquim José de França Júnior	1838	1883	<i>Cahio o ministério! Comedia original de costumes em tres actos</i>
Arthur Azevedo	1855	1892	<i>O tribofe: revista fluminense do ano de 1891</i>
Gastão M. Tojeiro	1880	1921/2017	<i>Onde canta o sabiá</i>
Gianfrancesco Guarnieri	1934	1957	<i>Eles não usam black-tie</i>
Miguel Falabella	1956	1988	<i>A partilha</i>
Paulo Sacaldassy	1966	2007	<i>Fulana Sicrana e Beltrana</i>
Paulo Gustavo A. M. de Barros	1978	2015	<i>Minha mãe é uma peça</i>

Fonte: Elaboração própria

Esta pesquisa se fundamenta no quadro teórico da Gramática Gerativa, que, conforme Chomsky (1994 [1986]),¹⁵ faz uma distinção entre língua-I e língua-E. A língua-I, ou língua internalizada, faz referência ao conhecimento do falante de sua língua materna, atingido por meio da experiência linguística durante a fase de aquisição da linguagem. Esta é a classificação de língua que é visada como objeto de estudo desse arcabouço teórico. A língua-E, ou língua externa, por sua vez, é um conceito de amplo alcance, pois engloba concepções de língua que não fazem referência direta ao sistema gramatical adquirido pelo falante e que são influenciadas por fatores externos, socio-históricos. É neste conceito onde estão a gramática normativa e a escrita. No entanto, para a sintaxe histórica, a língua escrita é frequentemente a principal fonte de dados e, assim, o trabalho diacrônico no quadro gerativista busca entender como as diferenças sintáticas observadas nos dados de língua-E refletem as mudanças de língua-I em gerações diferentes (Roberts, 2021). Portanto, ambos os critérios (ano de nascimento dos autores e ano de publicação dos textos) serão considerados para a análise, pois, apesar de o ano de nascimento estar mais próximo do período de aquisição da linguagem pelos autores, o trabalho é feito sobre dados de língua-E, que nem sempre apresentam apenas estruturas gramaticais adquiridas pelos falantes.

Na próxima seção, é feita uma síntese do que diziam as gramáticas normativas do século XIX sobre os pronomes interrogativos *que* e *o que*. A seguir, apresentam-se dados e, na sequência, discute-se como a norma pode se relacionar à variação nos autores de teatro brasileiro. Por fim, apresentam-se as considerações finais.

¹⁵ Dentro dos colchetes, apresenta-se o ano da publicação original; fora dos colchetes, referencia-se a edição consultada.

O que diz a norma

Levando em conta que o processo de gramaticalização de *o que* se deu em Portugal entre os séculos XVII e XVIII, pode-se esperar que, no século XIX, esse pronome já repercutisse nas gramáticas normativas em Portugal, bem como que a variação com *que* fosse percebida. A partir desse século, é possível verificar se o mesmo ocorria no Brasil. Considerando a hipótese de influência do PE na norma brasileira na segunda metade do século XIX, foram selecionadas algumas gramáticas publicadas nesse século – Sousa (1804) e Dias (1894) em Portugal; Grivet (1881), Júlio Ribeiro (1881), João Ribeiro (1889) no Brasil. Grivet (1881) era, na realidade, suíço, nascido em Villars-le-Terroir, e chegou ao Brasil em abril de 1856 (Almeida, 2022). A consulta de sua gramática é de interesse porque, por hipótese, sendo estrangeiro, pode destoar de gramáticos brasileiros contemporâneos a ele. Além disso, foi consultado também o *Compendio de grammatica da lingua nacional*, de Coruja (1835), para contemplar a primeira metade do século XIX no Brasil.

Sousa (1804), português, lista como pronomes interrogativos *que*, *qual* e *quem*. Esta gramática do início do século XIX não menciona *o que* como um pronome interrogativo, nem como um pronome relativo. O mesmo também foi observado no *Compendio de grammatica da lingua nacional* (Brasil), de Coruja (1835). Já Grivet (1881), em *Nova grammatica analytica da lingua portugueza*, publicada no Brasil, se refere a *que* como um pronome “indefinido”, diferenciando-o do pronome relativo, já que não há nenhuma referência. Para o gramático, *que* é um pronome “indefinido” quando significa *que coisa*, *coisa que* ou *que quantidade*, *que porção*. Grivet menciona a possibilidade de se colocar um artigo antes de *que*, dizendo que isso serve para “firmar” o sentido do pronome “indefinido” *que*, mas ressalta que *o que*, assim como *que*, é o mesmo que *que coisa/coisa que*, diferindo, portanto, de *o que* “demonstrativo”, que equivale a *aquilo que/aquele que*.

Julio Ribeiro (1881), brasileiro, contemporaneamente a Grivet (1881), prescreve que o artigo definido não deve aparecer antes dos pronomes “conjuntivos”¹⁶ quando estes forem “empregados interrogativamente”, dizendo que “« *O que queres?* – *O que te parece?* » e outras construcções identicas são incorrectas” (Ribeiro, 1881, p. 210, grifo do autor). Para embasar a sua prescrição, Ribeiro dá exemplos que aparecem em escritores “clássicos” – Camões, Vieira, Sousa Caldas, Bocage e Francisco Manuel. É no mínimo curioso que Grivet e Julio Ribeiro tenham posturas diferentes quanto ao uso de *o que*. Julio Ribeiro era brasileiro e falante nativo, enquanto Charles Grivet era suíço. O fato de que Grivet fosse estrangeiro, falante nativo de uma outra língua, talvez possa explicar o motivo de a presença do artigo antes de *que* não ser algo polêmico para ele, visto que o gramático se limita apenas a descrever a possibilidade de seu uso, sem julgá-la como

16 Quando menciona *pronomes conjuntivos*, o autor se refere ao que se entende atualmente por *pronomes relativos*: “[...] ‘que’, ‘quem’ referem-se sempre a um nome da clausula principal. Esse nome chama-se antecedente: pode ser masculino ou feminino; do singular ou plural. [...] Nas sentenças interrogativas o pronome ‘que’ admite depois de si o pronome a que se refere, ex. ‘Que homem é este?’ [...]” (Ribeiro, 1881, p. 226).

incorreta. Regressando a Portugal, vê-se na gramática de Dias (1894, p. 34) que o autor faz uma distinção entre os pronomes interrogativos adjetivos – *que*, *qual* e *quanto* – e os pronomes interrogativos substantivos – *que* (ou *o que*) e *quem*. O gramático menciona (como aqui) *o que* entre parênteses. Isso mostra que ele reconhece a variação, mas não a julga incorreta (como fez o gramático brasileiro Júlio Ribeiro alguns anos antes).

João Ribeiro (1889), gramático brasileiro, sugere que o uso de *o* (que considera um pronome) possa ser um “brasileirismo”: “A expressão *o que é a vida?* como a anteposição do pronome *o*, é provavelmente um brasileirismo.” (Ribeiro, 1889, p. 238, grifo do autor). O gramático justifica que “o uso clássico” não admite essa construção e que os “bons escritores” contemporâneos (no caso, do período em que publica essa gramática) demonstram isso. João Ribeiro ainda menciona que essa é a mesma sintaxe observada nas “línguas romanas”. Mais uma vez, um gramático brasileiro desencoraja o uso de *o* antes de *que* e ainda sugere que seja um “brasileirismo”. No entanto, em nota de rodapé, Ribeiro reconhece que há uso desse “pronome” em outras línguas românicas, como no galego (*o que*) e no florentino (*il che*).

Os dados

No *corpus* aqui utilizado, obtiveram-se os seguintes números para os dois autores do século XVIII:

Tabela 5. Variação entre autores nascidos no século XVIII em interrogativas diretas

País de nasc.	Autor (nasc.)	<i>Que</i> (nº)	<i>O que</i> (nº)	Freq. rel. de <i>o que</i> (aprox.)
Brasil	Silva (1705)	110	3	2,65%
Portugal	Costa (1757)	24	2	7,69%

Fonte: Elaboração própria

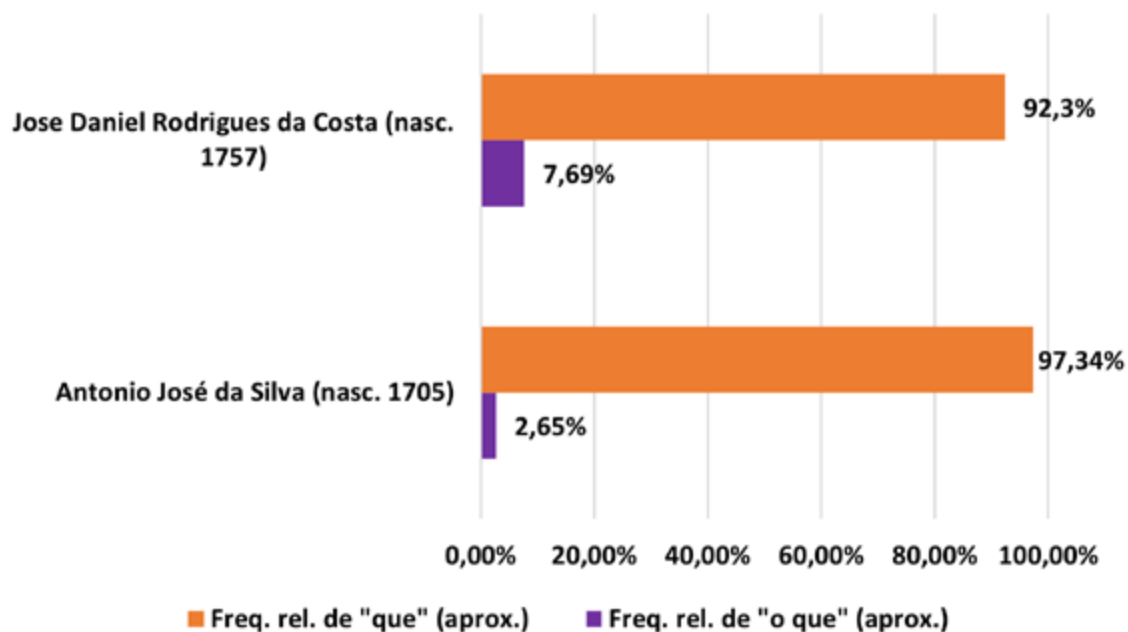
Tabela 6. Variação entre autores nascidos no século XVIII em interrogativas indiretas

País de nasc.	Autor (nasc.)	<i>Que</i> (nº)	<i>O que</i> (nº)	Freq. rel. de <i>o que</i> (aprox.)
Brasil	Silva (1705)	5	11	68,75%
Portugal	Costa (1757)	1	9	90%

Fonte: Elaboração própria

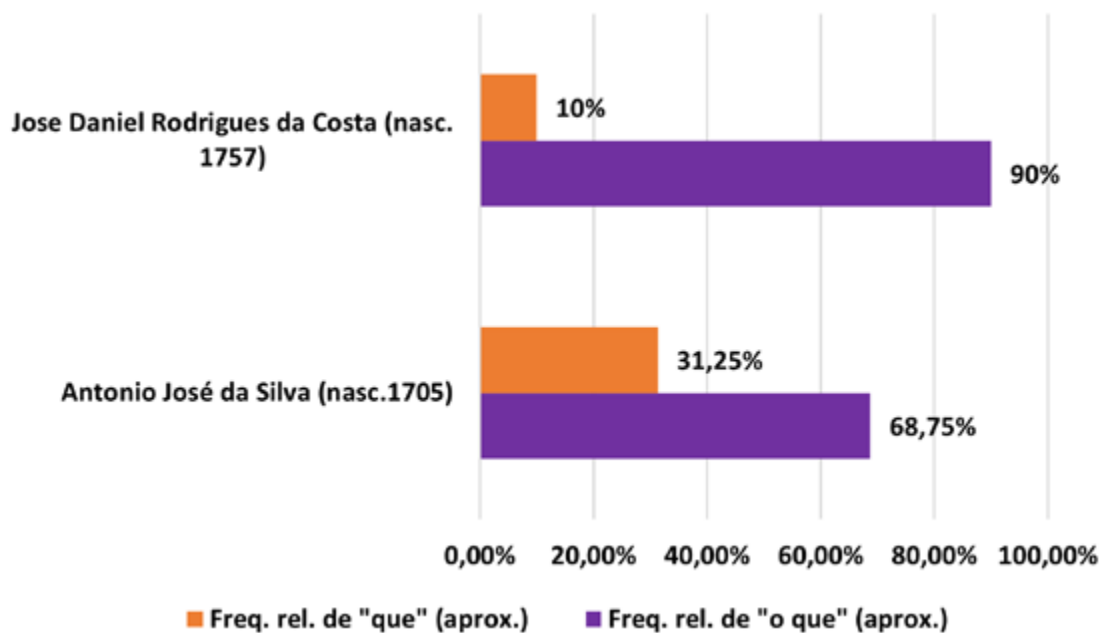
Como se nota, tanto o autor português como o luso-brasileiro se comportam de modo semelhante quanto ao uso de *que* e *o que* – sendo o primeiro preferido em interrogativas diretas e o último em interrogativas indiretas. Ilustra-se graficamente abaixo (Figura 1 e Figura 2).

Figura 1. Frequências relativas de *que* e *o que* nos autores nascidos no século XVIII em interrogativas diretas



Fonte: Elaboração própria

Figura 2. Frequências relativas de *que* e *o que* nos autores nascidos no século XVIII em interrogativas indiretas



Fonte: Elaboração própria

Para estes autores, não é possível realizar uma comparação ideal pelo ano de publicação das obras, porque o texto fonte de Antonio José da Silva é de 1759 e o de José Daniel Rodrigues da Costa é de 1973. Ainda que para a edição de 1973 o texto de base seja de 1879, ela está longe do século XVIII. A comparação por ano de nascimento, no entanto, mostra um comportamento semelhante entre os autores quanto ao uso de *que* e *o que*, fato que indica que, possivelmente, não houve alterações no texto original quanto ao uso dessas formas, já que há a questão da nacionalidade de Antonio José da Silva, que é, pode-se dizer, luso-brasileiro. O autor nasceu no Brasil, mas foi para Portugal aos sete anos de idade, após os pais terem recebido mandados de prisão da Inquisição, e frequentou a Universidade de Coimbra anos mais tarde (Silva, 2007). Não seria surpreendente, portanto, que os seus textos tivessem características em comum com os textos dos portugueses que não tiveram relação com o Brasil.

Nos autores brasileiros nascidos nos séculos XIX e XX, encontram-se os seguintes resultados quantitativos:

Tabela 7. Quantificação de *que* e *o que* em orações interrogativas diretas no teatro brasileiro dos autores nascidos nos séculos XIX e XX

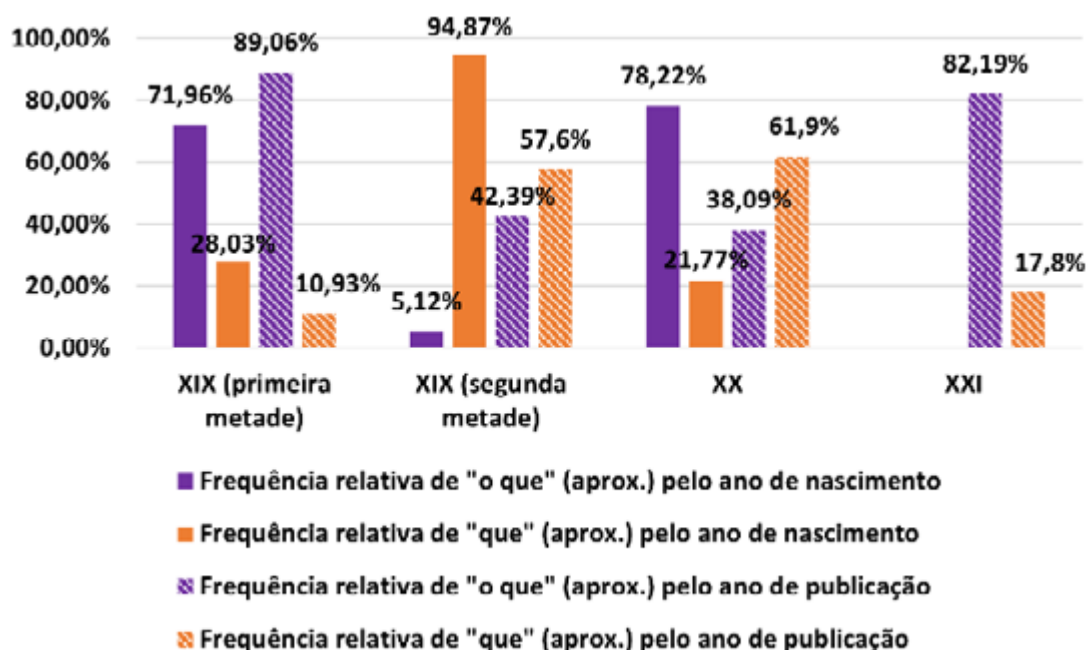
Nasc.	Publ. (fonte)	Autor	Que (nº)	O que (nº)	Freq. rel. de o que (aprox.)
1815	1838 (1871)	Pena	1	11	91,66%
1815	1845 (1853)	Pena	6	46	88,46%
1829	1858	Alencar	29	12	29,26%
1838	1883	França Jr.	1	26	96,29%
1855	1892	Azevedo	23	1	4,16%
1880	1921 (2017)	Tojeiro	51	3	5,55%
1934	1957	Guarnieri	12	20	62,5%
1956	1988	Falabella	2	17	89,47%
1966	2007	Sacaldassy	8	13	61,90%
1978	2015	Paulo Gustavo	5	47	90,38%

Fonte: Elaboração própria

Tomando como referência o ano de nascimento dos autores (primeira coluna), observa-se que, com exceção de José de Alencar, os autores nascidos na primeira metade do século XIX têm frequências relativas de *o que* mais altas em interrogativas diretas, superando a quantidade de *que*. Os dois autores nascidos na segunda metade, Arthur Azevedo e Gastão Tojeiro, apresentam baixa frequência relativa de *o que*, demonstrando preferência por *que* interrogativo. Para os autores nascidos no século XX, a frequência relativa de *o que* volta a subir. Já considerando o ano de publicação como referência, evidencia-se que

é mantida uma preferência por *o que* na primeira metade do século XIX;¹⁷ na segunda metade do século XIX, onde Alencar se encaixa por este critério, a mais frequente é *que*; no século XX, *que* ainda predomina (mas isto se deve sobretudo a Tojeiro, cuja publicação é da primeira metade do século). No século XXI, a preferência volta a ser *o que*. Desse modo, pode-se representar graficamente as frequências relativas de *que* e *o que* nesses períodos da seguinte forma:

Figura 3. Variação entre *que* e *o que* em interrogativas diretas no teatro brasileiro dos autores nascidos nos séculos XIX e XX



Fonte: Elaboração própria

Em orações interrogativas indiretas, que ocorrem em menor quantidade, observam-se os seguintes valores:

¹⁷ Está sendo utilizada como referência a data de publicação original para a análise. No caso, para a primeira metade do século XIX, trata-se dos textos de Pena – de 1838 e de 1845.

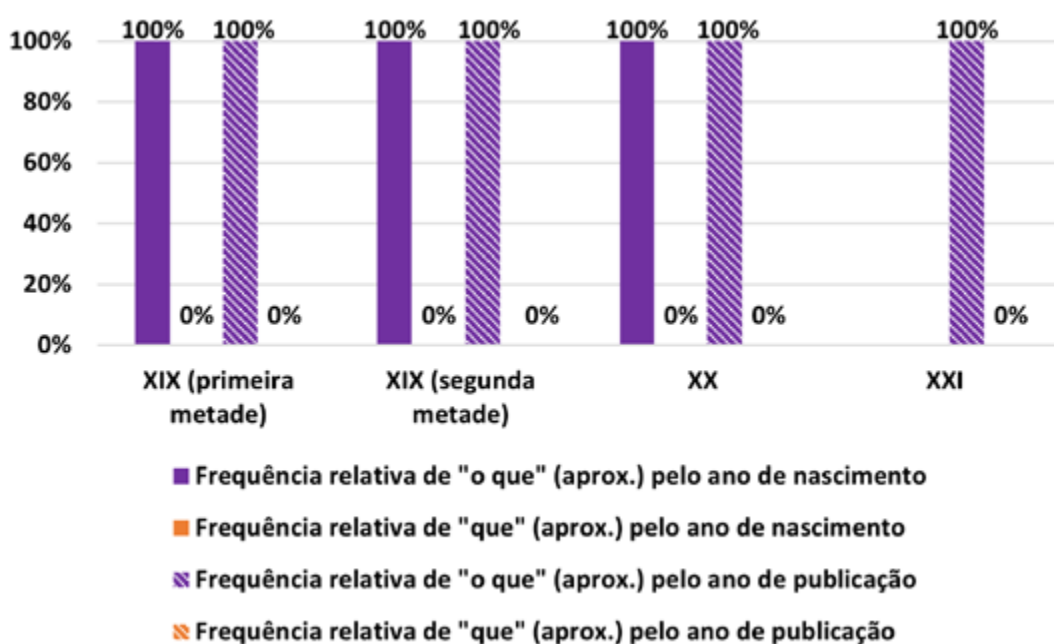
Tabela 8. Quantificação de *que* e *o que* em orações interrogativas indiretas no teatro brasileiro dos autores nascidos nos séculos XIX e XX

Nasc.	Publ. (fonte)	Autor	Que (nº)	O que (nº)	Freq. rel. de o que (aprox.)
1815	1838 (1871)	Pena	0	3	100%
1815	1845 (1853)	Pena	0	14	100%
1829	1858	Alencar	0	15	100%
1838	1883	França Jr.	0	12	100%
1855	1892	Azevedo	0	6	100%
1880	1921 (2017)	Tojeiro	0	11	100%
1934	1957	Guarnieri	0	5	100%
1956	1988	Falabella	0	10	100%
1966	2007	Sacaldassy	0	2	100%
1978	2015	Paulo Gustavo	0	24	100%

Fonte: Elaboração própria

Neste contexto, foram encontrados apenas dados com *o que*, de modo que a variação não está explícita (Figura 4).

Figura 4. Variação entre *que* e *o que* em interrogativas indiretas no teatro brasileiro dos autores nascidos nos séculos XIX e XX



Fonte: Elaboração própria

Nos autores nascidos no século XIX em Portugal, encontra-se a seguinte quantificação em interrogativas diretas:

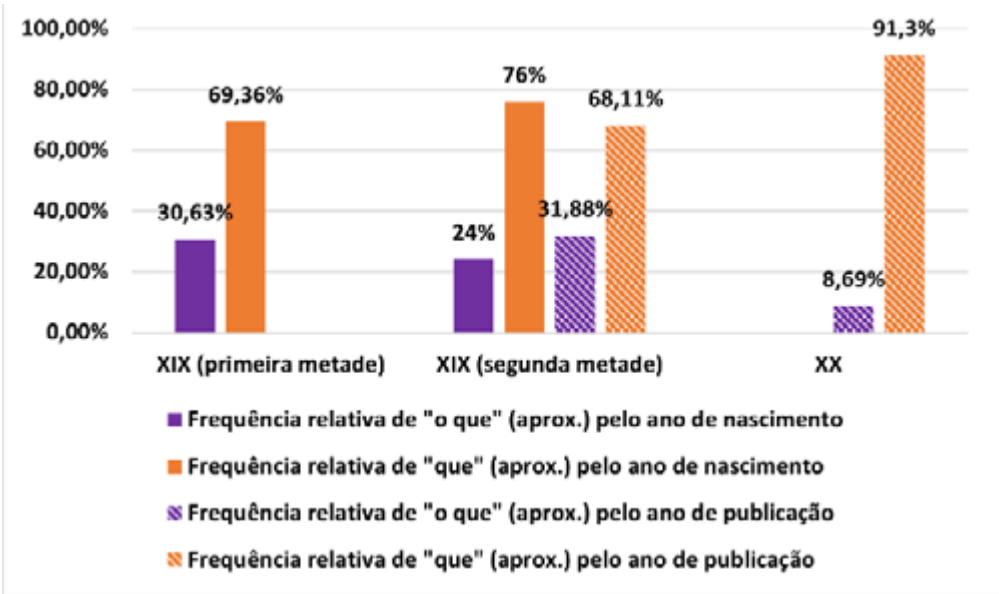
Tabela 9. Quantificação de *que* e *o que* em orações interrogativas diretas no teatro português dos autores nascidos no século XIX

Nasc.	Publ. (fonte)	Autor	<i>Que</i> (nº)	<i>O que</i> (nº)	Freq. rel. de <i>o que</i> (aprox.)
1815	1899	Cascais	29	3	9,37%
1825	1871	Branco	31	0	0%
1842	1869	Chagas	17	31	64,58%
1856	1897	Mesquita	17	10	37,03%
1881	1916	Brun	21	2	8,69%

Fonte: Elaboração própria

Nota-se que apenas Chagas tem uma porcentagem de *o que* maior que 50%. Há, portanto, neste *corpus* português, uma preferência por *que* em interrogativas diretas. Pelo critério do ano de nascimento, observa-se que, tanto na primeira como na segunda metade do século XIX, a porcentagem *que* é mais elevada (Figura 5). Para o *corpus* português, as publicações são, em maioria, da segunda metade do século XIX, havendo apenas André Brun que muda para o século XX. Assim, não é possível comparar totalmente os dados em função desse critério de análise. No entanto, observa-se que a preferência se mantém por *que* na segunda metade do século XIX e que André Brun, com publicação no século XX, usa mais *que* (Figura 5).

Figura 5. Frequências relativas de *que* e *o que* em interrogativas diretas no teatro português dos autores nascidos no século XIX



Fonte: Elaboração própria

Nas interrogativas indiretas, obtiveram-se os seguintes números:

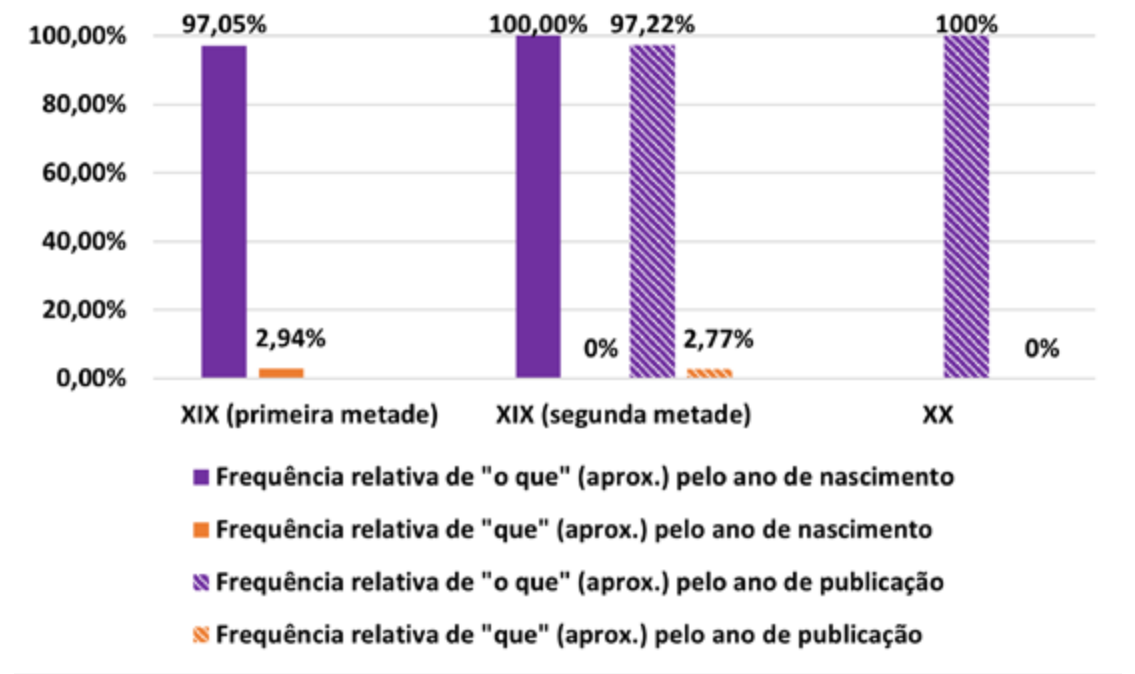
Tabela 10. Quantificação de *que* e *o que* em orações interrogativas indiretas no teatro português dos autores nascidos no século XIX

Nasc.	Publ. (fonte)	Autor	<i>Que</i> (nº)	<i>O que</i> (nº)	Freq. Rel. de <i>o que</i> (aprox.)
1815	1899	Cascais	1	8	88,88%
1825	1871	Branco	0	13	100%
1842	1869	Chagas	0	12	100%
1856	1897	Mesquita	0	2	100%
1881	1916	Brun	0	9	100%

Fonte: Elaboração própria

Assim como ocorreu com o PB, as porcentagens de *o que* são altas. Observa-se, portanto, uma ampla preferência por *o que* em interrogativas indiretas durante todo o século XIX, ilustrada graficamente na Figura 6. Pelo critério do ano de publicação, não há alterações relevantes, exceto o fato de não haver dados da primeira metade e a mudança de André Brun, que passa para o século XX e que não tem nenhum dado com *que*.

Figura 6. Frequências relativas de *que* e *o que* em interrogativas indiretas no teatro português dos autores nascidos no século XIX



Fonte: Elaboração própria

A norma *versus* os dados

Conforme exposto na seção *O que diz a norma*, observou-se que os gramáticos brasileiros Julio Ribeiro e João Ribeiro, que tiveram gramáticas publicadas na segunda metade do século XIX, são conservadores quanto ao uso de *o que*, condenando o seu uso. João Ribeiro até chega a dizer se tratar de um “brasileirismo”. Como mostrado por Galves, Andrade e Silva (2026, no prelo), em textos portugueses, *o que* já aparece em interrogativas diretas desde o século XVIII. Esse conservadorismo é coerente com os resultados apresentados na seção anterior, em que se nota nos autores brasileiros de teatro uma postura mais conservadora com relação ao uso de *o que* em orações interrogativas diretas na segunda metade do século XIX. Posteriormente, no século XX, pelo critério do ano de nascimento, o uso de *o que* volta a aumentar em interrogativas diretas. No entanto, pelo critério do ano de publicação, *que* ainda predomina, o que se deve, sobretudo, ao alto número de ocorrências dessa forma em Tojeiro, cuja publicação data ainda da primeira metade do século XX e é o único desse grupo que nasceu na segunda metade do século XIX. Por esse critério, França Jr. também se distancia das publicações da segunda metade do século XIX, já que faz mais uso de *o que*.

O texto de José de Alencar é outro caso que, a depender do critério de análise escolhido, pode ou não estar destoante dos outros de seu período. Se considerarmos o ano de nascimento, o autor se esquadra na primeira metade do século XIX, tendo mais ocorrências de *que*, enquanto os outros autores de sua geração no *corpus* apresentam mais ocorrências de *o que*. Isto pode tanto aproximar Alencar do português clássico (PC) (cf. Tabela 1) como do PE que, no *corpus* aqui utilizado, apresenta mais dados de *que* no século XIX. Para fazer uma afirmação ou outra de forma decisiva, seria necessário que mais aspectos gramaticais fossem analisados (e em mais obras do autor). No entanto, em Schei (2003), que estuda a posição dos clíticos em obras literárias brasileiras, é possível observar, na quantificação dos dados, que José de Alencar¹⁸ estava dentre os autores que variavam, isto é, que tinha características do PB e do PE: apresentava ocorrências de próclise (ainda que poucas) em posição inicial; tinha alta taxa de ênclise em contexto de oração coordenada; apesar de usar mais ênclise do que próclise em contexto de verbo precedido de sujeito, tinha também uma das porcentagens mais altas de próclise (dentre os autores do seu *corpus*) neste contexto; em orações subordinadas, Alencar usa mais próclise do que ênclise mas, do *corpus* da autora, ele é o que mais usa ênclise neste contexto. Esses resultados de Schei para Alencar sugerem uma tendência de adoção da gramática do PE, de modo que até “escape” um desvio dessa norma que não pode ser associado a uma gramática brasileira, a ênclise em orações subordinadas, mas sem apagar totalmente características do PB (como a próclise em posição inicial).

Aliás, características do PB não passaram despercebidas em Portugal na recepção de *Iracema*, na segunda metade do século XIX. Na época, o autor e crítico português

18 A obra do autor analisada por Schei (2003) foi o romance *Lucíola*.

Manuel Pinheiro Chagas (que faz parte do *corpus* português aqui analisado) criticou o distanciamento da norma portuguesa por parte do autor brasileiro, o que, ao ver dele, impossibilitava o reconhecimento da autonomia literária do Brasil (Moreira, 2013). Isto é relevante porque, como se observou, Chagas é o autor português com mais ocorrências de *o que* em interrogativas diretas e é o único que usa mais essa forma do que a outra (cf. Tabela 9). Possivelmente, essa variação não era tão saliente em Portugal como aparentemente foi no Brasil do século XIX, vide os dados e as prescrições das gramáticas brasileiras. Note-se que, se for considerado o ano de publicação como referência para a análise, Alencar se enquadra na segunda metade do século XIX e, assim, não é surpreendente que haja mais ocorrências de *que*. Leite (1996), ao abordar a questão da criação de neologismos por Alencar, explica que o autor tinha um discurso contraditório, pois sempre se preocupou “em negar a atitude purista ortodoxa”, mas sua atitude “não era anti-ortodoxa porque estava sempre ancorada nos clássicos ou na tradição lusitana [...]” (Leite, 1996, p. 108). Segundo a autora, o que influenciava a postura de Alencar era resultado de duas tendências da época: a força da norma lusitana vs. a força do nacionalismo típico do Romantismo. Leite também recupera uma síntese de Pimentel Pinto (s.d.) acerca do uso de artigo por Alencar: “Elimina-se o artigo antes dos nomes próprios, salvo quando não se conhece o indivíduo, ou se pode confundi-lo com outro, e, também antes de nomes de familiares; Usa-se o artigo com sobrenomes e com certos nomes geográficos” (Pimentel Pinto, s.d. *apud* Leite, 1996, p. 106-107). Portanto, Alencar adota um uso mais restritivo dos artigos. Por hipótese, também seria possível pensar que o autor monitorasse mais o uso de *o que* em interrogativas diretas, ainda que deixasse algumas ocorrências passarem.

Nas gramáticas portuguesas do século XIX consultadas, apenas se encontra menção a *o que* em Dias (1894), que não condena o uso da forma e apenas a apresenta como uma variante de *que*, o que sugere uma postura não prescritiva. Os dados do teatro português mostram um avanço relativamente tímido de *o que* durante o século XIX, uma vez que a maior parte dos dados são de ocorrências de *que* (cf. Tabela 9 e Figura 5). Pode-se dizer que, no caso do Brasil, as gramáticas indicam uma pressão normativa na segunda metade do século XIX para um padrão do PE, favorecendo o uso de *que*, forma que mais aparece nos textos portugueses do século XIX. Esse efeito da norma na segunda metade do século XIX no Brasil, voltada para o PE, também já foi atestado por Pagotto (1992, 2013, 2018) e Carneiro (2005) a respeito da colocação de clíticos. Pagotto (2013), em particular, ao investigar a posição dos clíticos nas constituições do Império (de 1824) e da República (de 1891) – portanto em textos formais –, observa que a primeira se volta para a gramática do PC/ e a última para a gramática do PE. Para os dados analisados aqui, não se pode afirmar que haja, na primeira metade do século XIX, uma adequação ao PC/ no Brasil, pois *o que* é a forma inovadora e aparece mais nos textos da primeira metade do século do que a forma *que*, mais antiga e presente em interrogativas diretas no PC/ (cf. Tabela 1). O que parece estar evidente é que as peças brasileiras seguem um padrão: ou da primeira ou da segunda metade do século XIX, havendo alguns casos que destoam a depender do critério utilizado – ano de nascimento do autor ou ano de publicação.

Por fim, retoma-se que, no que toca ao século XVIII, o autor português José Daniel Rodrigues da Costa e o autor luso-brasileiro Antonio José da Silva não apresentam diferenças significativas quanto ao uso de *que* e *o que*. Quanto às orações interrogativas indiretas, nota-se que já há uma preferência por *o que* nesses casos, fato possivelmente relacionado à ambiguidade entre orações interrogativas indiretas e orações relativas livres (contexto sintático onde *o que* também aparece) na história da língua (Silva, 2023; Galves; Andrade; Silva, 2026, no prelo).

Considerações finais

Buscou-se, neste trabalho, apresentar e discutir a dinâmica de variação diacrônica entre *o que* e *que* no PB em comparação com o PE, a partir de um *corpus* de teatro de autores brasileiros e portugueses nascidos entre os séculos XVIII e XX. Ao mesmo tempo em que se procurou chamar a atenção para um assunto ainda pouco discutido no PB,¹⁹ este trabalho se debruçou sobre a forte tensão entre língua-I e língua-E existente em dados diacrônicos do PB, em especial no século XIX, quando se observa uma diferença entre autores da primeira e da segunda metade do século. Na variação analisada, o aumento de *que* na segunda metade do século XIX em interrogativas diretas não parece estar relacionado ao processo de aquisição, portanto, à língua-I, mas a uma pressão normativa orientada para o PE. Essa tendência de uso de uma gramática mais próxima do PE na segunda metade do século XIX, já observada em Pagotto (1996, 2013, 2018) e Carneiro (2005), se alinha com os dados aqui analisados. Tratando-se de uma variação e considerando que *o que* parece já ter sido incorporado pelo PB na primeira metade do século XIX (cf. Figura 3), pode-se reforçar a ideia de que haveria uma competição de gramáticas (Kroch, 1994) no Brasil envolvendo três sistemas – o PC/, o PE e o PB (Carneiro, 2005; Carneiro; Galves, 2010), sendo o uso do PE conservador e relacionado à pressão normativa.

Agradecimentos

O presente trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processo nº 2024/08111-0. Também recebeu apoio anterior da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Parte dos dados analisados foi coletada com apoio anterior da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Brasil. Processo nº 2021/01980-5. Agradeço também à Prof.^a Dr.^a Charlotte Galves pelas orientações no desenvolvimento deste trabalho e aos pareceristas pelos apontamentos. Eventuais erros são de minha responsabilidade.

19 Cabe mencionar Medeiros Junior (2014), que, em sua tese sobre as relativas livres no PB, discute a diacronia de *o que* e propõe um percurso de mudança de $[_{DP} \text{ o } [_{CP} \text{ que}]]$ para $[_{DP} \text{ o que}]$.

Referências

- ALMEIDA, A. B. F. de. *A sintaxe na Nova Grammatica Analytica da Lingua Portuguesa (1881), de Charles Adrien Olivier Grivet (1816-1876): uma análise historiográfica*. 2022. Dissertação (Mestrado em Linguística) – Universidade Federal da Paraíba, João Pessoa, 2022. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/24054>. Acesso em: 4 jul. 2024.
- CARNEIRO, Z. *Cartas brasileiras (1809-1904): um estudo lingüístico-filológico*. 2005. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2005. Disponível em: <https://hdl.handle.net/20.500.12733/1603052>. Acesso em: 14 jul. 2024.
- CARNEIRO, Z.; GALVES, C. Variação e Gramática: colocação de clíticos na história do português brasileiro. *Revista de Estudos da Linguagem*, [S. l.], v. 18, n. 2, p. 7-38, 2010. DOI: <https://doi.org/10.17851/2237-2083.18.2.7-38>
- CHOMSKY, N. *O conhecimento da língua: sua natureza, origem e uso*. Lisboa: Caminho, 1994.
- CORUJA, A. A. P. *Compendio de grammica da lingua nacional*. Porto Alegre: Tipografia V.F. de Andrade, 1835. Disponível em: <https://digital.bbm.usp.br/handle/bbm/7895>. Acesso em: 11 set. 2025.
- DIAS, A. E. da S. *Grammatica portugueza elementar*. 9. ed. rev. Lisboa: A. Ferreira Machado, 1894. Disponível em: <https://purl.pt/29127>. Acesso em 4 jul. 2024.
- GALVES, C.; ANDRADE, A.; SILVA, V. M. A variação entre o que e que na história do português: um estudo baseado em corpora. In: SANTOS, I.; INVERNO, L.; CARVALHO, M. J. (ed.). *Pelo tempo e pelo espaço. Estudos em Homenagem a Clarinda de Azevedo Maia*. Coimbra, Portugal: Imprensa da Universidade de Coimbra, no prelo.
- 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. *Journal of Historical Syntax*, v. 93, n. 3, p. 152-180, set. 2017. DOI: <https://doi.org/10.18148/hs/2017.v0i0.28>.
- GRIVET, C. A. O. *Nova grammica analytica da lingua portugueza*. Rio de Janeiro: Typ. De G. Leuzinger & Filhos, 1881. Disponível em: <https://digital.bbm.usp.br/handle/bbm/7507>. Acesso em: 4 jul. 2024.

KROCH, A. Morphosyntactic Variation. In: BEALS, K. et al. (ed.). *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*. [s.l.] Chicago Linguistic Society, 1994.

LEITE, M. Q. *O Purismo Lingüístico: suas manifestações no Brasil*. 1996. Tese (Doutorado em Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 1996. DOI: <https://doi.org/10.11606/T.8.1996.tde-19092022-111951>.

LOPES-ROSSI, M. G. Estudo diacrônico sobre as interrogativas no português do Brasil. In: ROBERTS, I.; KATO, M. A. *Português brasileiro: uma viagem diacrônica*. São Paulo: Contexto, 2018. p. 241-266.

MATEUS, M. H. M. et al. *Gramática da língua portuguesa*. 7. ed. Lisboa: Caminho, 2006.

MEDEIROS JUNIOR, P. *Orações relativas livres do PB: sintaxe, semântica e diacronia*. 2014. Tese (Doutorado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2014. DOI: <https://doi.org/10.47749/T/UNICAMP.2014.939421>.

MOREIRA, M. E. José de Alencar e a crítica portuguesa do século XIX. *Convergência Lusíada*, [S. l.], v. 24, n. 29, p. 195-203, jun. 2013. Disponível em: <https://www.convergencialusiada.com.br/rcl/article/view/196>. Acesso em: 12 set. 2025.

PAGOTTO, E. G. Clíticos, mudança e seleção natural. In: ROBERTS, I.; KATO, M. A. (org.) *Português brasileiro: uma viagem diacrônica*. 3. ed. São Paulo: Contexto, 2018. p. 143-160.

PAGOTTO, E. G. A Norma das Constituições e a Constituição da Norma no século XIX. *Revista LETRA*, v. 1, p. 31-50, 2013.

PAGOTTO, E. G. *A posição dos clíticos em português: um estudo diacrônico*. 1992. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1992. DOI: <https://doi.org/10.47749/T/UNICAMP.1992.46498>.

RIBEIRO, J. *Grammatica portugueza*. São Paulo: Typ. De Jorge Seckler, 1881. Disponível em: <https://digital.bbm.usp.br/handle/bbm/2097>. Acesso em 4 jul. 2024.

RIBEIRO, J. *Grammatica Portugueza Exame de Portuguez* 3º ano. Rio de Janeiro: Livraria Classica de Alves e C., 1889.

ROBERTS, I. *Diachronic syntax*. 2. ed. Oxford, Reino Unido; New York, NY: Oxford University Press, 2021.

SCHEI, A. A colocação pronominal na literatura brasileira do século XIX. *Filologia e Linguística Portuguesa*, [S. l.], n. 5, p. 57–84, 2003. DOI: 10.11606/issn.2176-9419.v0i5p57-84. Disponível em: <https://revistas.usp.br/flp/article/view/59700>. Acesso em: 3 fev. 2026.

SILVA, A. J. da. *As comédias de Antonio José, o Judeu: Vida de d. Quixote, Vida de Esopo, Anfitrião e Guerras do alecrim*. São Paulo: Martins Fontes, 2007.

SILVA, V. M. *A ambiguidade entre orações interrogativas indiretas e orações relativas livres: um estudo diacrônico sob o viés do interrogativo o que*. 2023. Monografia (Bacharelado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2023. Monografia defendida, não publicada oficialmente. Disponível em: <https://osf.io/2ezh9/>. Acesso em: 30 set. 2025.

SOUSA, M. D. de. *Grammatica portugueza*. Coimbra: Real Imp. da Universidade, 1804. Disponível em: <https://purl.pt/17363>. Acesso em 4 jul. 2024.