

Emanoel Cesar Pires de Assis<sup>1</sup> 

*Universidade Estadual do Maranhão*

*emanoel.uema@gmail.com*

Daniel Lopes<sup>2</sup> 

*Universidade Estadual do Maranhão*

*daniell139@gmail.com*

## **A estatística textual computadorizada e a literatura brasileira: uma análise do romance *Miragem*, de Coelho Neto**

### **Resumo:**

Neste artigo, revelamos, por meio de uma leitura estilométrica, características do estilo de escrita de Coelho Neto em *Miragem* (1895). Ademais, comparamos o estilo do escritor com o estilo de três escritores contemporâneos seus: Aluísio Azevedo, *O Cortiço* (1890); Adolfo Caminha, *Bom-Crioulo* (1895); e Machado de Assis, *Dom Casmurro* (1900). A metodologia de investigação adotada nesse trabalho tem como suporte a estilometria (Brandão, 2006; Cúrcio, 2006, 2007, 2013; Freitas, 2007; Paiva, 2013; Assis, 2013; e Silva, 2014) e dialoga, sempre que possível, com a crítica literária direcionada a Coelho Neto e sua obra, principalmente no tocante ao seu estilo. Dentre os nossos achados, comprovamos que os usos de adjetivos, verbos, e advérbios em –mente apontados como “excessivos” pela crítica por parte de Coelho Neto na composição do seu estilo são, aparentemente, normais quando comparamos o seu estilo com o estilo de seus contemporâneos.

---

<sup>1</sup> Doutor em Letras (UFSC). Professor do Departamento de Letras e do Mestrado em Letras da UEMA.

<sup>2</sup> Graduado em Letras pelo CESC/UEMA, membro do Grupo de Pesquisas em Literatura, Arte e Mídias- LAMID/CNPq- UEMA.

**Palavras-chave:** literatura brasileira, Coelho Neto, estatística textual, estilometria

**Abstract:**

**The Computerizes Textual Statistics and the Brazilian Literature: An Analysis of the Romance *Miragem*, by Coelho Neto**

In this article, we reveal, through a stylometric reading, characteristics of Coelho Neto's style in *Miragem* (1895). In addition, we compare the style of the writer with the style of three contemporary writers: Aluísio Azevedo, *O Cortiço* (1890); Adolfo Caminha, *Bom-Crioulo* (1895); and Machado de Assis, *Dom Casmurro* (1900). The research methodology adopted in this study is supported by stylometry (Brandão, 2006, Cúrcio, 2006, 2007, 2013, Freitas, 2007, Paiva, 2013, Assis, 2013 and Silva, 2014) and dialogues, whenever possible, with the literary criticism directed to Coelho Neto and his work, mainly in relation to his style. Among our findings, we have found that the uses of adjectives, verbs, and adverbs in Coelho Neto's criticism of the composition of his style are apparently normal when compared to the style of other writers.

**Keywords:** Brazilian literature, Coelho Neto, textual statistics, stylometry

### **Coelho Neto: um escritor de ambígua figuração**

Henrique Maximiano “Coelho Neto” é um escritor de “ambígua figuração” dentro do ideário da literatura brasileira. Em sua época, fora considerado o escritor mais lido do Brasil, hoje é visto como um escritor controvertido, de uma produção excessiva, que perpassa praticamente todas as correntes estéticas da literatura brasileira, do fim do século XIX a meados do século XX, e encontra-se ainda, infelizmente, no limbo literário, procurando, através de seus volumes e gêneros escritos, espaço para ser inserido ou, ao menos, ocupar uma posição melhor nas Histórias da Literatura Brasileira.

Com o propósito de entendermos melhor essa emblemática figuração, resolvemos realizar um estudo de caráter estilístico pelo viés da “leitura estilométrica” em uma de suas obras, o romance *Miragem* (1895).

A crítica, ao tocar de maneira qualitativa e avaliar, especificamente, o plano formal dos gêneros escritos por Coelho Neto, costumeiramente

o acusa de ser prolixo e “afetado” por usar, em excesso, classes gramaticais como verbos, adjetivos, e advérbios (Bosi, 2013). Tal prática, do ponto de vista da crítica, sobrecarrega o estilo do autor, dificultando as raras leituras de suas obras pelo público leitor.

A falta de um estudo mais acurado em relação a este mesmo estilo fez das acusações formuladas e exploradas “o cavalo de batalha de mil condenações, às vezes levianas, às vezes ridículas” em torno do nosso escritor em tela (Farias, 1963: 11). Destarte, entendemos que uma das várias pertinências desse estudo estilométrico é começarmos a preencher a lacuna que ainda se encontra aberta a respeito do estilo desse escritor maranhense.

A metodologia de investigação denominada de estilometria é, em pleno curso do século XXI, no Brasil, pouco utilizada. Entretanto, de maneira gradativa, trabalhos da mesma natureza que o nosso (Brandão, 2006; Freitas, 2007; Paiva, 2013; e Assis, 2013) têm ajudado a torná-la bastante promissora. Atuando como metodologia de investigação científica no campo literário, a estilometria surge a partir do interesse de áreas distintas e que se complementam mutuamente, dentre elas, principalmente, a área da literatura, da linguística, da informática e da estatística, caracterizando-a assim como uma metodologia interdisciplinar.

Dessa forma, procuraremos obter, viabilizados pelo software Hyperbase, de Etienne Brunet, dados estilométricos retirados do romance *Miragem* e traçar as características do estilo de escrita de Coelho Neto, cotejando o seu estilo com o de três escritores contemporâneos seus: Aluísio Azevedo (*O Cortiço*); Adolfo Caminha (*O Bom Crioulo*); e Machado de Assis (*Dom Casmurro*).

### **Procedimentos para um estudo estilométrico: preparação dos textos**

Para criar a nossa base textual, com os textos dos quatro escritores, tivemos que seguir alguns procedimentos para que os textos pudessem ser devidamente inseridos no programa Hyperbase e, assim, retirar deles os dados estilométricos que nos interessariam para as análises

e interpretações. Etapas como: coleta dos textos; revisão e adequação da grafia foram devidamente realizadas para a criação da base textual. Nessas etapas são feitas também a limpeza textual, comparação editorial e atualização dos textos para serem submetidos ao Hyperbase.

A padronização gráfica permite um melhor tratamento dos textos e evita erros que possam estar ligados a edições diferentes das obras com grafia do português igualmente diferente. A partir disso, é possível obter resultados estatísticos que são gerados pelo programa e que o pesquisador precisa, juntamente com o conhecimento da leitura da obra, interpretar.

Ao todo, nosso corpus de textos conta com 312.316 ocorrências (N) e 23. 815 formas/vocábulo (V). Assis (2013) entende por “forma” qualquer unidade que se encontra no texto, ou seja, o conceito abarca tanto as palavras quanto os sinais de pontuação, já as “ocorrências” são todas as unidades do texto, bem como as formas e suas repetições. A base textual, então, tem as seguintes características:

<b>TEXTOS</b>	<b>OCORRÊNCIAS</b>	<b>FORMAS</b>	<b>ABREVIACÃO</b>
O Cortiço	101.040	11.203	CORTI
Bom-crioulo	47.614	7.096	CRIOL
Miragem	81.019	11.072	MIRA
Dom Casmurro	82.643	8.688	DOMC

Tabela 1 – Número de ocorrências e vocábulos nos romances. Fonte: os autores.

Como podemos perceber, os textos têm extensões diferentes. Logo, o cálculo estatístico dos dados que desejamos investigar não acontecerá a partir da quantidade de vezes ou da média em que determinada forma aparece, o que pode gerar dados equivocados, na medida em que a proporção das palavras e outros elementos textuais, ou em nosso caso, as formas investigadas adjetivos, verbos e advérbios estão intrinsecamente relacionadas com a extensão de cada texto. Dessa maneira, visando obter resultados que nos dão maior confiabilidade, a análise

de cada dado estatístico dos textos investigados ocorrerá a partir do cálculo do desvio padrão em relação ao corpus.

### **Riqueza lexical de Coelho Neto em *Miragem***

A respeito das críticas em direção ao estilo de Coelho Neto, Arnaldo Niskier, em Coelho Neto e a Modernidade (2010), enfatiza e comenta a campanha modernista opressora sobre o escritor maranhense. Niskier (2010) destaca que “o modernismo condenou-o como representante do passadismo, acusado de afetação, palavreado rebuscado e enfático, abuso de termos incomuns, prolixidade e helenismo”. Ademais, diz ainda que “a vitória do modernismo se fez como se houvesse necessidade de abater um grande inimigo, no caso, Coelho Neto”.

Instigados em querer saber se Coelho Neto possui, de fato, um dos mais ricos vocabulários em língua portuguesa, conforme sugerido por seu filho, resolvemos investigar a sua riqueza lexical a fim de percebermos se tal observação sobre o seu estilo pode ser verificada.

Verônica Cúrcio explica que o cálculo da riqueza lexical em textos literários, na estilometria, baseia-se na relação da medida estatística entre “o número total de palavras repetidas e diferentes de um mesmo texto e o número de palavras que o compõem” (2013: 84). Ou seja, dentro dos princípios dos estudos estilométricos, refere-se à razão entre o número de palavras diferentes (vocábulos ou formas) e o número total de palavras (ocorrências).

Entendida assim, podemos deduzir que um determinado escritor terá o vocabulário rico e variado a partir de um número maior de vocábulos novos que aparecem por meio de seu estilo de escrita no(s) seu(s) texto(s), sem a possibilidade de muita repetição, do contrário, mais repetitivo e restrito em vocabulário será o texto, o que indicará uma possível pobreza lexical. Cúrcio (2013: 84) resume essa noção de riqueza lexical da seguinte maneira: “quanto maior o número de vocábulos novos, maior será a riqueza e a variedade do vocabulário”.

Com o auxílio do software Hyperbase, podemos obter por meio de cálculos estatísticos baseados em distribuições de frequências e na extensão relativa do vocabulário dos textos a riqueza lexical dos autores

em duas perspectivas: vocabulário geral ou o índice de hapax (palavras de ocorrência única). Sendo assim, para uma primeira apreciação dos resultados da riqueza lexical e hapax dos estilos dos nossos escritores em seus textos, vejamos a figura abaixo:

The screenshot shows a software window titled "Riqueza do vocabulário e hapax". It contains two tables of data. The top table lists authors with columns for rank (n°), total words (total), unique words (hapax), and other metrics. The bottom table shows a similar set of data for a different group of authors.

	n°	total	hapax	total	hapax	total	hapax	total	hapax
1	3208	3289	-188	-14.80	3874	-6.86	COELHO		
2	2986	3093	-107	-11.19	3889	5.77	CRUZ		
3	3372	3374	-2	-1.50	3364	17.66	MIRA		
4	3029	3410	-381	-13.85	2698	-11.72	SOMM		
Total	28816				11995				

	n°	total	hapax	total	hapax	total	hapax	total	hapax
1	218	889	-124	-14.30	1499	-9.20	COELHO		
2	3045	3777	-482	-9.88	1187	8.76	CRUZ		
3	2864	2617	-247	-3.13	3021	19.00	MIRA		
4	3321	2281	-1040	-15.17	3249	-15.27	SOMM		
Total	14619				6187				

Figura 1 – Riqueza do vocabulário e hapax nos romances. Fonte: Hyperbase.

No conjunto geral dos textos, a partir dos resultados estilométricos expostos acima, podemos verificar que das formas totais (23. 815 vocábulos), 11. 995 são hapax. Ou seja, quase metade do vocabulário presente nos textos dos nossos escritores em tela não se repete, o que em termos estilísticos demonstra uma habilidade considerável dos nossos prosadores na composição dos seus estilos em seus romances. Entretanto, ao verificarmos separadamente os dados estilométricos expostos na figura 1, percebemos que há uma preponderância da riqueza lexical e dos números de hapax no estilo de Coelho Neto, que é superior ao estilo de seus contemporâneos. Vejamos novamente os resultados da figura 1, na tabela abaixo, por posição de riqueza lexical.

<b>Posição de R. L</b>	<b>Nº de Vocábulos</b>	<b>Nº de Hapax</b>	<b>Romances</b>
1º	11. 072	3. 954	<b>MIRA</b>
2º	11. 203	3. 574	<b>CORTI</b>
3º	8. 688	2. 608	<b>DOMC</b>
4º	7. 096	1. 859	<b>CRIOLO</b>

Tabela 2 - Posição de riqueza lexical e hapax por romances. Fonte: os autores.

A tabela acima nos dá a primeira comparação entre o estilo de Coelho Neto e o estilo de seus contemporâneos. Nessa comparação, com esses dados primários, podemos constatar que Coelho Neto, em *Miragem*, tem um quantitativo de diversidade lexical superior aos seus contemporâneos, seu texto apresenta quase quatro mil (4.000) hapax.

Podemos afirmar, a partir dos dados acima, que o romance *Miragem* tem aspectos positivos no tocante ao léxico, ou seja, é um romance que dá sinais e traz, em sua composição estilística, vocabulários variados, sem a possibilidade de muita repetição, o que em termos estilísticos demonstra a habilidade de Coelho Neto com o uso das palavras. Tal amostra, igualmente, ratifica a consideração da crítica sobre o autor quando o denomina de “mago da palavra”.

### **Ocorrências de adjetivos, verbos e advérbios em *Miragem***

De acordo com Murari (2011: 26), desde o início da profusa carreira de Coelho Neto como escritor, “foram frequentes as admoestações ao autor acerca de seu estilo superabundante e imaginoso, da excentricidade de sua prosa oblíqua, seduzida [...] pelo fraseado sinuoso, pelo descritivismo detalhista e extenuante”. De tais admoestações negativas, temos este comentário de Luís Murrat direcionado aos talentos iniciais de Coelho Neto: “faz uns folhetins aos sábados. Tem talento, mas abusa do adjetivo e tem a mania do Oriente” (Lima, 1958: 34, itálico nosso). Quanto ao uso “demasiado” dos “verbos” e ainda dos “adjetivos” em seu estilo, Jomar Moraes, citado em Martins (2014: 32), profere que Coelho Neto fora o “malabarista do verbo” e o “libidinoso da adjetivação” dada a tamanha importância ao uso dos dicionários.

Vejamos se o que diz a crítica pode ser comprovado a partir de uma leitura estilométrica.

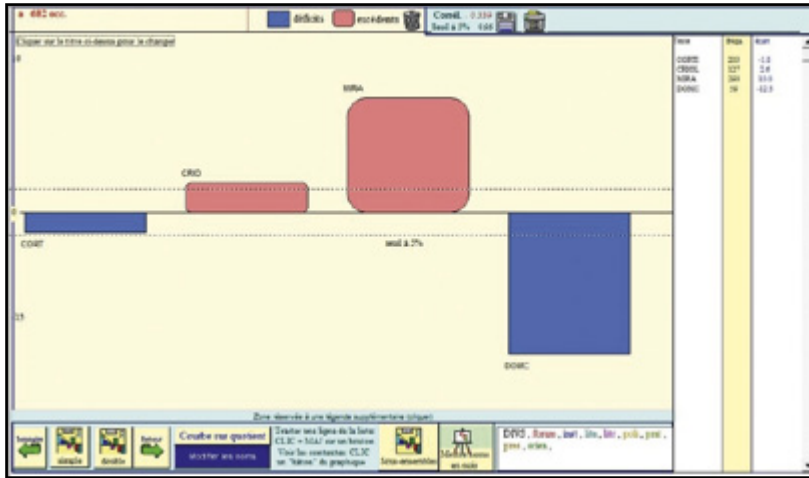


Gráfico 3 – Ocorrências de Adjetivos nos romances. Fonte: Hyperbase.

Os dados apontam exatamente aquilo que Bosi (2013) comenta sobre o estilo de Coelho Neto quanto a essa classe gramatical analisada, ou seja, que o escritor de *Miragem* escora o seu estilo em classes como o adjetivo. Acusa também outra observação de Mendes e Ignácio (2010: 2) quando destacam o “anseio pela adjetivação numerosa, rara e erudita” de Coelho Neto em sua prosa. Destarte, podemos confirmar que, realmente, em comparação com os romances dos seus contemporâneos, há uma maior frequência de adjetivos no romance *Miragem*, tornando assim os discursos da crítica literária numericamente acertados quanto ao uso de adjetivos no estilo do autor, ao menos no romance analisado.

Entretanto, como bem observado por Muralli (2011) ao dizer que a crítica condenava o escritor Coelho Neto pelo descritivismo detalhista e extenuante em suas obras, podemos entender que, para uns, pode ser considerado um problema estilístico, para outros, assegura uma qualificação do enredo, dos personagens e da própria narrativa como um



todo. Dessa maneira, é inegável as considerações da crítica ao dizerem que o escritor maranhense tem um apreço especial pelos adjetivos.

É válido mencionar outra observação importante quanto ao uso de adjetivos pelo autor Coelho Neto em seus romances, a que foi feita por Martins (2014). Ao comparar dois outros romances do escritor, Turbilhão e A Conquista, a pesquisadora percebe que, em número de adjetivos, A Conquista contém mais adjetivos e hapax que Turbilhão. Perante os nossos dados estatísticos e a observação de Martins (2014), no tocante aos adjetivos, é pertinente o seguinte questionamento: Coelho Neto oscila o uso dessa classe gramatical em seus romances?

Com o intuito de verificar, agora sobre o uso de advérbios, o que a crítica fala sobre o estilo de Coelho Neto, também verificamos, por meio do Hyperbase, como essa classe gramatical se comporta no corpus criado e quais são os textos que se destacam quanto ao seu uso.

Assim, obtivemos o seguinte gráfico:

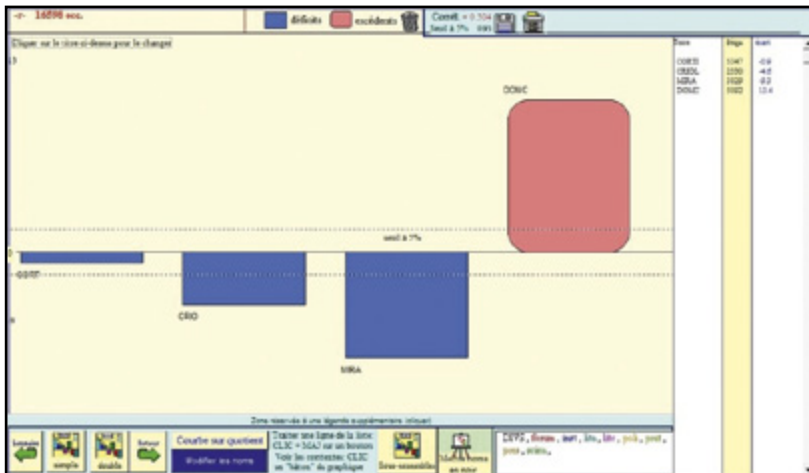


Gráfico 4 – Ocorrências de Advérbios nos romances. Fonte: Hyperbase.

Os dados obtidos revelam que não é exatamente Coelho Neto quem faz uso em demasia de advérbios, pelo menos não se levamos em conta os romances em tela, mas, antes, Machado de Assis. Uma vez

que Dom Casmurro sobressai quanto ao item analisado. O que, de certa maneira, se coloca contrário àquilo que Bosi (2013) afirma ao dizer que Coelho Neto, continua o “velho estilo” de José de Alencar, escorando-se em classes como o advérbio. Dado muito semelhante é exposto por Martins (2014) quando afirma que a obra de Coelho Neto não se caracteriza pelo uso excessivo de advérbios. Assim, nossa pesquisa, aliada à de Martins (2014), reitera que o estilo de Coelho Neto não pode ser assim classificado.

A análise nos revela que Dom Casmurro, como já mostrado no outro gráfico, apresenta predominância de advérbios, bem como de subordinação, numerais e preposições seguidas de pronomes. Enquanto em *Miragem* há mais preposições, interjeições e preposições seguidas de determinantes.

### **Considerações finais**

A proposta do nosso estudo estilométrico foi buscar breves características do estilo de escrita de Coelho Neto possíveis de serem detectadas a partir de análises estatísticas quanti-qualitativas, cotejando o seu estilo com o estilo de alguns de seus contemporâneos, e procurando confirmar, refutar, ou ampliar as “certezas” da crítica literária sobre o estilo do autor maranhense quando diz ser caracterizado por uma riqueza lexical e usos constantes de adjetivos, verbos e advérbios.

Utilizando o Hyperbase, comprovamos, quantitativamente e qualitativamente, que a riqueza lexical do estilo de Coelho Neto, encontrada no romance *Miragem*, é superior quando comparada com as riquezas lexicais dos estilos de outros contemporâneos a ele. Ratificamos no corpus, igualmente, que o uso da classe de adjetivos aparece com mais frequência no estilo do escritor maranhense do que no estilo dos demais autores confrontados. Em relação a essa classe de palavras que, embora vista, pela crítica, como um aspecto negativo, no estilo do autor de *Miragem*, conseguimos encontrar respaldo e explicações plausíveis à luz da teoria do estilo para entender o porquê de Coelho Neto primar tanto pelo uso de adjetivos em seus enredos.

Quanto ao uso de verbos e advérbios no corpus, a nossa análise provou que Dom Casmurro, em muito, ultrapassa *Miragem* quando o quesito verificado é o uso dos verbos e dos advérbios. E nem por isso o escritor de Dom Casmurro foi tão criticado e perseguido ao usar essas classes gramaticais em seu estilo quanto o escritor de *Miragem*.

As informações retiradas da nossa análise podem, de alguma maneira, lançar por terra ou pôr em dúvida a crítica negativa voltada a Coelho Neto quanto ao uso das classes gramaticais analisadas em seu estilo. Elas também podem assegurar as afirmações de Farias (1963) e Broca (1981) ao dizerem que Coelho Neto foi vítima de negligência, juízos críticos sumários, levianos e falhos por parte da crítica, principalmente no tocante ao seu estilo.

### Referências bibliográficas

- ASSIS, E., “Atribuição de autoria utilizando análises estatísticas: uma experiência com a relação abreviada”, [on-line] <https://www.periodicos.ufsc.br/index.php/textodigital/article/view/1807-9288.2013v9n1p24> – 13.09.2017.
- BOSI, A. (2013), *História concisa da literatura brasileira*, Cultrix, São Paulo.
- BROCA, B. (1981), *Ensaio da Mão Canhestra: Cervantes, Goethe, Dostoiévski, Alencar, Coelho Neto, Pompéia*, Polis, São Paulo; INL, Brasília.
- CÚRCIO, V. (2013), *Palavras de Rosa: análise estilométrica da obra de João Guimarães Rosa*, UFSC, Florianópolis.
- FARIA, O. de (1963), *Coelho Neto: romance*, Livraria Agir Editora, Rio de Janeiro.
- MENDES, R. F. C; IGNÁCIO, E., “Entre dois modus vivendi: arcaísmo e modernidade em *Turbilhão*, de Coelho Neto”, [on-line] [http://www.prp2.ueg.br/sic2010/apresentacao/trabalhos/pdf/linguistica/seminario/entre\\_dois\\_modus\\_vivendi.pdf](http://www.prp2.ueg.br/sic2010/apresentacao/trabalhos/pdf/linguistica/seminario/entre_dois_modus_vivendi.pdf) - 20.09.2017.
- MAYDANA, C. J. D. (2010), *Decifrando os enigmas da modernidade em esfinge, de Coelho Neto*, FURG, Rio Grande.
- MARTINS, F. M. (2014), *Compilação, anotação e análise linguístico-computacional do corpus Coelho Neto, um corpus de textos literários dos séc. XIX e XX*, UFC, Fortaleza.

- MURARI, L., “Sob o tênue véu da ficção: três eventos da história brasileira”, [on-line] <http://revistaseletronicas.pucrs.br/ojs/index.php/navegacoes/article/view/9435> – 31.08.2017.
- NETO, P. C. (1958), “Imagem de uma vida” em: Aguilar, L. (org.), Coelho Neto. Obra Seleta Vol. I: Romances, J. Aguilar, Rio de Janeiro, pp. 83-110.
- NISKIER, A., “Coelho Neto e a modernidade” [on-line] <http://www.academia.org.br/artigos/coelho-neto-e-modernidade> – 30.10.2017.
- PAIVA, D. M. (2013), *Um poeta particular: estudo estilométrico da poesia de H. Dobal*, UFPI, Teresina.
- LIMA, H. (1958), “Coelho Neto: as duas faces do espelho” em: Aguilar, L. (org.), Coelho Neto. Obra Seleta Vol. I: Romances, J. Aguilar, Rio de Janeiro, pp. 11-82.
- ROSSO, M. (2010) *Lima Barreto versus Coelho Neto: um Fla-Flu literário*, DIFEL, Rio de Janeiro.
- SILVA, L. G. (2014), *As visões de amor na poesia de Augusto dos Anjos: estudo estilométrico*, UFPI, Teresina.