

V Encontro Internacional de Língua Portuguesa e Relações Lusófonas



LUSOCONF 2023

LIVRO DE ATAS **Proceedings**

Editores:

Carla Sofia Araújo
Ana Paula Monte
Paula Odete Fernandes
João Sérgio Sousa
Vitor Gonçalves

Instituto Politécnico de Bragança
julho de 2024

Ficha Técnica

Título

LUSOCONF2023

V Encontro Internacional de Língua Portuguesa e Relações Lusófonas: livro de atas

Editores

Carla Sofia Araújo

Instituto Politécnico de Bragança, Portugal

Ana Paula Monte

Instituto Politécnico de Bragança, Portugal

Paula Odete Fernandes

Instituto Politécnico de Bragança, Portugal

João Sérgio Sousa

Instituto Politécnico de Bragança, Portugal

Vitor Gonçalves

Instituto Politécnico de Bragança, Portugal

Capa

António Meireles e Vitor Gonçalves

Instituto Politécnico de Bragança, Portugal

Suporte técnico

Clarisse Pais, Serviços de Documentação e Bibliotecas, Instituto Politécnico de Bragança, Portugal

Edição

Instituto Politécnico de Bragança

Campus de Santa Apolónia

5300-253 Bragança

Portugal

Data de edição: julho de 2024

ISBN: 978-972-745-328-3

DOI: 10.34620/978-972-745-328-3

Handle: <http://hdl.handle.net/10198/25534>

URL: www.lusoconf.ipb.pt

Email: lusoconf@ipb.pt

Estudo lexical da poesia de Eugénio de Andrade

Lexical study of the poetry of Eugénio de Andrade

Carla Sofia Araújo [0000-0002-5318-0960]

carla.araujo@ipb.pt

Instituto Politécnico de Bragança, Portugal

Centro de Estudos em Letras, Universidade de Trás-os-Montes e Alto Douro, Portugal

Resumo. Este estudo orienta-se por fundamentos metodológicos da Linguística de *Corpus*. A Linguística de Corpus pode ser utilizada como um importante auxiliar de observação e análise de textos. No caso do texto literário, a Linguística de Corpus configura uma abordagem que fornece dados linguísticos e informações sobre os textos a que outros tipos de métodos de análise textual não automática não permitem aceder. Neste trabalho, a partir do programa computacional Nooj (Silberztein, 2003), efetuamos uma análise lexicométrica, baseada na análise estatística das palavras-tema, tendo em vista a delimitação de possíveis campos temáticos na poesia de Eugénio de Andrade.

Palavras-Chave: Linguística de Corpus, Nooj, Palavras-tema, Campos temáticos.

Abstract. This study is guided by the methodological foundations of Corpus Linguistics. Corpus Linguistics can be used as an important aid in the observation and analysis of texts. In the case of literary texts, Corpus Linguistics configures an approach that provides linguistic data and information about texts to which other types of textual analysis methods do not allow automatic access. In this work, using the computational program Nooj (Silberztein, 2003), we generate a lexicometric analysis, based on the statistical analysis of the theme-words, with a view to delimiting possible thematic fields in the poetry of Eugénio de Andrade.

Keywords: Corpus linguistics, Nooj, Theme words, Theme fields.

1 Introdução

Este artigo apresenta um estudo lexical da poesia de Eugénio de Andrade, aplicando os critérios e os princípios da Linguística de Corpus e usando o programa computacional Nooj (Silberztein, 2003). A partir das ferramentas do Nooj, será possível analisar a lista de *tokens* e identificar as formas mais usadas pelo autor, tendo em vista delimitar campos temáticos na poesia de Eugénio de Andrade.

Este artigo é constituído por duas partes. Na primeira parte, apresentamos as potencialidades da Linguística de Corpus para o estudo do texto literário. Na segunda parte, recorrendo ao Nooj, apresentaremos o estudo lexical da poesia de Eugénio de Andrade, partindo dos *tokens* mais frequentes na obra. Primeiramente, apresentamos os dados gerais do corpus. A seguir, organizamos uma listagem de palavras-tema, a partir da listagem dos *tokens* com maior frequência, fornecida automaticamente pelo Nooj. O estudo lexical termina com a apresentação dos campos temáticos, delimitados a partir das palavras-tema presentes na poesia de Eugénio de Andrade.

2 Linguística de Corpus e estudo do texto literário

Como nos lembra Brait (2000), existe uma falsa dicotomia entre língua e literatura, uma vez que a língua e a literatura estão intrinsecamente ligadas.

O estudo da língua em uso permite uma abordagem integrada do sistema linguístico subjacente à comunicação. Tal sistema, que é para muitos designado por lexicogramática (Biber; Conrad; Reppen, 1998; Berber Sardinha, 2020), articula a vertente lexical e gramatical de maneira reciprocamente dependente. A Linguística de Corpus permite-nos contabilizar o número de ocorrências das formas

linguísticas de um corpus, tendo em vista análises multivariadas, como, por exemplo, a análise multidimensional (Berber Sardinha & Veirano Pinto, 2019).

A Linguística de Corpus possibilita a análise da frequência dos itens linguísticos, proporcionando-nos identificar associações lexicais, mercê da força da atração que as palavras exercem umas sobre as outras, tal como Firth (1957, p. 11) explicou com a sua famosa frase: «you shall judge a word by the company it keeps», concluindo-se que a linguagem se baseia em padrões léxico-gramaticais, sabidos e partilhados pelos falantes.

As ferramentas computacionais da Linguística de Corpus permitem-nos também, entre outros aspetos, identificar os pacotes lexicais (“lexical bundles”) (Biber, Conrad & Cortes (2004), palavras-chave (Scott, 1997), as formas exclusivas ou formas de frequência 1 e as palavras-tema, como veremos na secção dedicada ao estudo lexical da poesia de Eugénio de Andrade, através do programa Nooj (Silberstein, 2003).

Para além do Nooj, existem outros programas disponíveis, como, por exemplo, o programa WordSmith Tools (Scott, 2022), muito utilizado para a comparação de frequências, o AntConc, o ProtAnt e o Graphcoll (Anthony, 2022; 2023; Brezina, 2016), que permitem verificar todas as ocorrências dos itens linguísticos, a partir da disponibilização de listas de palavras, concordâncias e linhas de concordância, apresentando e organizando todos os dados linguísticos observados.

A Linguística de Corpus exerce uma notável influência na pesquisa linguística em vários centros internacionais. Os métodos de pesquisa baseados em corpus têm sido usados para estudar a literatura, tendo-se tornado cada vez mais populares (Biber, 2011). As suas potencialidades têm também sido exploradas no campo dos estudos literários, estando disponíveis diversas publicações que apresentam aplicações de conceitos e metodologias da Linguística de Corpus para a Literatura (Hoey *et al.*, 2007). Semino e Short (2004) preconizam a adoção de uma perspetiva de cooperação para a utilização dos métodos da Linguística de Corpus e para o estudo do texto literário, salientando a utilidade da combinação da metodologia baseada em corpus com métodos baseados na intuição. Ao analisarem as conexões entre a Linguística de Corpus e a Estilística Literária, Hoey *et al.* (2007) destacam que a Estilística concentra-se em como dizemos o que dizemos e a Linguística de Corpus também defende que o que dizemos está dependente da forma, ou seja, dos padrões que são confirmados nos corpora. Os referidos autores sublinham também o facto de o foco da Linguística de Corpus e da Estilística Literária ser distinto, uma vez que a Estilística se concentra no que faz um texto ser distinto. Além disso, esta disciplina estuda desvios das normas linguísticas que provocam efeitos artísticos e representam formas criativas de uso da linguagem. Já a Linguística de Corpus dirige a sua atenção principalmente para os usos repetidos e típicos, focando-se nos padrões repetidos.

Leech e Short (1981) sublinham que “não há nenhuma técnica infalível para selecionar o que é significativo. Temos de estar conscientes do efeito artístico total para cada texto e das particularidades linguísticas que encaixam nesse todo”.

Decorrente da aplicação da metodologia da Linguística de Corpus, os resultados poderão ser semelhantes aos das pesquisas ancoradas em análises tradicionais. De facto, mercê do uso que o autor faz da linguagem, a abordagem da Linguística de Corpus pode convergir ou divergir das impressões intuitivas do leitor (Zyngier *et al.*, 2011). Contudo, ainda que os dados quantificados confirmem aquilo que já se conhece, tal situação não constitui um constrangimento, mas antes uma confirmação de que a abordagem é fidedigna. Por conseguinte, a Linguística de Corpus pode fornecer ao crítico literário as ferramentas e os dados necessários para que a literatura possa ser analisada a partir de uma nova abordagem, que é mais objetiva do que a abordagem tradicional, de base hermenêutica.

Na secção a seguir, apresenta-se o estudo lexical da poesia de Eugénio de Andrade.

3 Estudo lexical da poesia de Eugénio de Andrade

Através da análise linguística efetuada pelo Nooj, podemos aceder aos dados gerais caracterizadores do texto, apresentados na tabela 1.

Tabela 1: Dados Gerais do Corpus – “Poesia”.

Dados Gerais do <i>Corpus</i> – “Poesia”	
Unidades de texto	12056

N.º de caracteres	350175 (273784 letras; 64207 espaços em branco; 11553 outros delimitadores)
Tokens	76720
Word forms	64536
Delimiters	11553
Anotações	204012
Ambiguidade	4859 tipos diferentes de ambiguidade
Unidades linguísticas não ambíguas	2876

Fonte: Elaboração própria.

Automaticamente, o programa computacional Nooj apresenta-nos os *tokens* e as respetivas frequências. Para o objetivo do nosso trabalho, interessa-nos a listagem dos *tokens* apresentada por ordem decrescente da sua frequência.

A análise das formas mais frequentes permite-nos verificar que as formas mais frequentes são palavras funcionais ou gramaticais. Especificamente, as cinco formas mais frequentes são as seguintes: “A/a”, “O/o”, “De/de”, “E/e”, com uma frequência de 2702, 2404, 1980, 1396, 1359, respetivamente.

Seguidamente, analisamos a listagem dos *tokens* e selecionamos os 60 *tokens* mais frequentes, tendo em vista elaborar a listagem de palavras-tema e das respetivas frequências, que apresentamos de seguida. Para a elaboração da listagem dos 60 *tokens* mais frequentes, seguimos como critério a seleção de nomes, de adjetivos e de verbos principais (Mateus *et al.* 2003).

Tabela 2: Palavras-tema de “Poesia”, de Eugénio de Andrade.

Frequência	TOKEN	Frequência	TOKEN	Frequência	TOKEN	Frequência	TOKEN
303	Luz/luz	142	Mãos/mãos	99	Águas/águas	69	Lume/lume
290	Só/só	133	Vento/vento	97	Música/música	68	Cal/cal
249	Corpo/corpo	124	Mão/mão	96	Rumor/rumor	67	Flor/flor
243	Olhos/olhos	118	Coração/coração	95	Morte/morte	67	
							Palavra/palavra
219	Água/água	117	Amor/amor	87	Dias/dias	67	Neve/neve
207	Sol/sol	117	Rio/rio	87	Nome/Nome	65	Mãe/mãe
202	Verão/verão	110	Dia/dia	84	Fim/fim	63	ardor
199	Noite/noite	110	Tempo/tempo	81	Alegria/alegria	63	céu
190	Mar/mar	109	Palavras/palavras	78	Rosto/rosto	61	Vezes/vezes
175	Terra/terra	109	Pedra/pedra	78	Veze/veze	60	Desejo/desejo
165	Boca/boca	106	Olhar/olhar	78	ar	60	
							Homem/homem
157	Casa/casa	106	Sei/sei	76	Tarde/tarde	60	Lugar/lugar
157	Ser/ser	102	Azul/azul	75	Chuva/chuva	60	Voz/voz
156	Silêncio/silênci o	100	Lábios/lábios	74	Morrer/morrer	57	Folhas/folhas
148	Sombra/sombra	98	Manhã/manhã	69	Branco/branco	55	dedos

Fonte: Elaboração própria.

Quando iniciamos a análise das palavras-tema, verificamos fenómenos de ambiguidade potencial, que integra a maior parte das formas linguísticas, dado que a listagem dos *tokens* fornecida pelo Nooj apresenta as palavras-tema sem os contextos de ocorrência. Por conseguinte, procedemos à extração de concordâncias no Nooj, tendo em vista detetar fenómenos de homonímia, parcial ou absoluta (Chaves, 2013). Quando as formas lexicais possuem iguais formas gráficas e fonéticas, sentidos independentes e pertencem a classes de palavras diferentes, estamos perante casos de homonímia parcial. Estamos perante casos de homonímia absoluta em palavras pertencentes à mesma classe e portadoras de sentidos independentes. Deste modo, procedemos à análise de todos os contextos de ocorrência de cada uma das palavras-tema do corpus e consultamos as definições das palavras-temas no dicionário (<https://www.infopedia.pt/>), atendendo a que a análise das concordâncias e a consulta do dicionário asseguram uma análise mais objetiva.

Como podemos constatar na figura 1, a palavra-tema “rio”, que configura um fenómeno de homonímia parcial, ou seja, “rio” [verbo rir] e “rio” [nome], conforme se verifica na consulta do dicionário, remete-nos para a forma do verbo “rir” e para o nome masculino singular “rio”.

Através da leitura das concordâncias da palavra-tema “rio”, constata-se que a forma “rio” corresponde ao nome masculino singular de “rio” e à primeira das três aceções do nome “rio” (“1. Curso natural de água que nasce, em geral, nas montanhas e vai desaguar ao mar, a um lago ou a outro rio, ou, por vezes, se entranha na terra; 2. figurado aquilo que corre como este curso de água; 3. figurada quantidade considerável; abundância”). Deste modo, como se pode verificar na tabela 3, a palavra-tema “rio” insere-se no campo temático de “Natureza”.

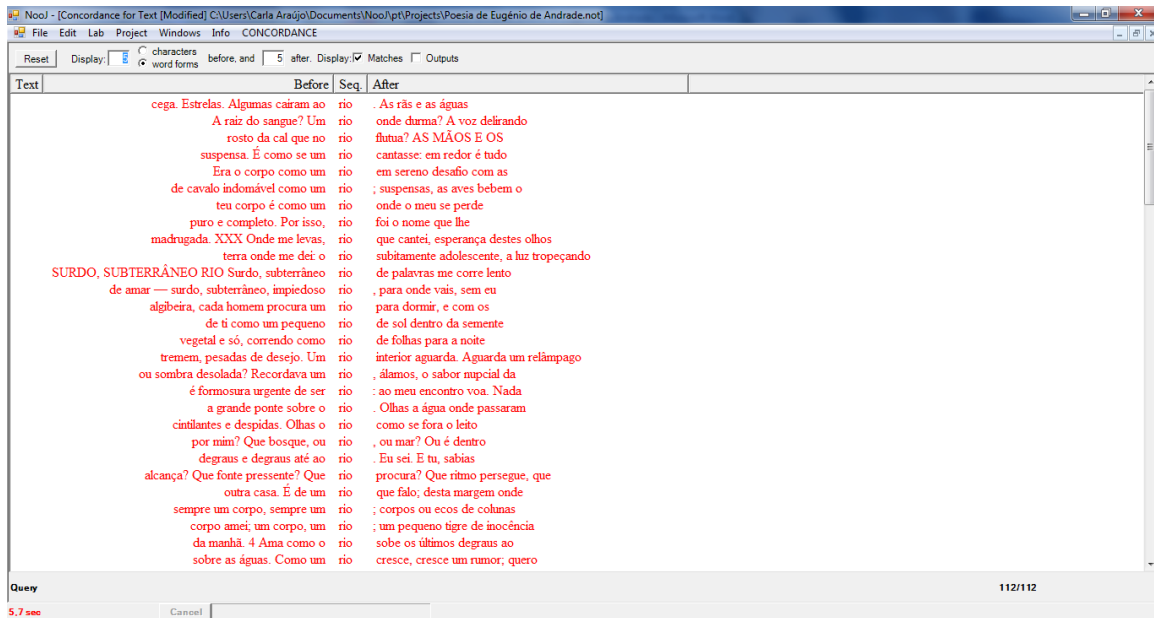


Figura 1: Concordâncias da palavra-tema “rio”.
Fonte: Nooj.

Como podemos constatar na figura 2, a palavra-tema “verão”, que também configura um fenómeno de homonímia parcial, ou seja, “verão” [verbo ver] e “verão” [nome], conforme se verifica na consulta do dicionário, remete-nos para o futuro do indicativo do verbo “ver” e para o nome masculino singular “verão”.

Através da leitura das concordâncias da palavra-tema “verão”, constata-se que a forma “verão” corresponde ao nome masculino singular de “verão”, por isso, como se pode verificar na tabela 3, a palavra-tema “verão” insere-se no campo temático de “temporalidade”.



Figura 2: Concordâncias da palavra-tema “verão”.
Fonte: Nooj.

Já a palavra-tema “sobre” não consta da listagem de palavras-tema, uma vez que, em todas as ocorrências, estamos perante a preposição “sobre”. Como se pode verificar na figura 3, não se trata nem do nome masculino nem do imperativo ou do presente do conjuntivo do verbo sobrar.

Text	Before	Seq.	After
	sobe à torre abandonada e	sobre	a azenha parada tomba desamparado
	do que lágrimas, coisas frias,	sobre	as tuas mãos abandonadas à
	E no teu rosto aberto	sobre	o mar cada palavra era
	princípio. Como se houvesse nuvens	sobre	nuvens, e sobre as nuvens
	houvesse nuvens sobre nuvens, e	sobre	as nuvens mar perfeito, ou
	esconde para amar outra luz	sobre	as areias. Um pedaço de
	encontro vem a grande ponte	sobre	o rio. Olhas a água
	lume. Para orvalho te queria	sobre	as horas transtornadas. Para a
	colunas, rios ou súbitas janelas	sobre	dunas; corpos: docéis, doirados montes
	rosto no meu. Que faz	sobre	o teu peito! No teu
	súbito o céu se fecha	sobre	dunas e barcos, e cada
	devagar escorre a última luz	sobre	as areias. Que diremos ainda
	pelas dunas. Soltei os olhos	sobre	aquele corpo, o coração latindo
	dezaesseis potros brancos na colina	sobre	as águas. Como um rio
	-me cada passo, recomeço, pedra	sobre	pedra, a juntar palavras; quero
	junta. Música das primeiras chuvas	sobre	o feno. Só aroma. Abelha
	AS JANELAS As janelas abrem	sobre	as fontes. Abrem para o
	boca mais vulnerável. A marcar	sobre	os teus flancos o itinerário
	assim ardente. Pousas as mãos	sobre	o meu rosto, e vais
	machucados do silêncio, vertente aguda	sobre	o mar — já as águas
	pedra cintilante dos teus joelhos,	sobre	o teu corpo o meu
	só lembrados de um verão	sobre	outros lábios. RUPTURA A noite
	outono, desatado beijo a beijo	sobre	a palha; ficam as nuvens
	espessa ondulação do vento é	sobre	o corpo a própria exaltação
	própria nudez erramos na noite	sobre	os membros o peso obscurecido
	água era o meu corpo	sobre	as noites havia outras colinas
	apenas dormisse com os lódiãos	sobre	o coração. AS FONTES As
	beber no escuro. Deita-te	sobre	o meu peito inclina até

Figura 3: Concordâncias da palavra-tema “sobre”.

Fonte: Nooj.

A palavra-tema “são” remete-nos para mais um fenómeno de homonímia parcial. Nesse sentido, a leitura das concordâncias permitiu-nos verificar que a forma “são” nem é o adjetivo nem o nome “são”, já que, em todas as ocorrências, esta forma diz respeito ao presente do indicativo do verbo ser, como se pode verifica na figura 4.

Text	Before	Seq.	After
	ACORDE Onde passou o vento	são	altas as ervas, e os
	rasos de água. Só elas	são	estrelas penduradas nos meus dedos
	abrir — é um cheiro forte,	são	, de frutos naturais que a
	as palavras que te digo	são	duras, mãe, e o nosso
	essas mãos que nem sequer	são	flores que se dêem: abertas
	flores que se dêem: abertas	são	apenas abandono, fechadas são pálpebras
	abertas são apenas abandono, fechadas	são	pálpebras imensas carregadas de sono
	tu dizias: os teus olhos	são	peixes verdes. E eu acreditava
	eram realmente peixes verdes. Hoje	são	apenas os meus olhos. É
	As palavras que te envio	são	interditas até, meu amor, pelo
	Um pássaro e um navio	são	a mesma coisa quando te
	do mundo. Alguns pensam que	são	as mãos de deus — eu
	de deus — eu sei que	são	as mãos de um homem
	penetram, indiferentemente. Não lhes toquem:	são	amor e bondade. Mais ainda
	MADRIGAL Os anjos que prometes	são	apenas o rosto triste dos
	rua. Os seus pés descalços	são	formosos. São formosos e leves
	de alegria e de terror;	são	a grande razão, a única
	pelas mãos, que já não	são	mais que fugidios ecos de
	estremecem. Desamparadas, inocentes, leves. Tecidas	são	de luz e são a
	Tecidas são de luz e	são	a noite. E mesmo pálidas
	ó barco onde as bandeiras	são	todas de alegria, água súbita
	No teu ombro respiro. Belos	são	os navios, altos, estreitos. Feliz
	No teu ombro respiro. Belas	são	as areias fúlvias de verão
	silêncio: em abril os dias	são	frágeis, impacientes e amargos; os
	Onde estão as barcas? Onde	são	as ilhas? ESCRITO NO MURO
	Ariadne ou não, os olhos	são	azuis. Azuis de um azul
	gregos de Ariadne, que não	são	de Ariadne nem são gregos
	não são de Ariadne nem	são	gregos, desses olhos que se

Figura 4: Concordâncias da palavra-tema “são”.

Fonte: Nooj.

De igual modo, a palavra-tema “entre” configura um fenómeno de homonímia parcial, isto é, “entre” [verbo entrar] e “entre” [preposição]. Esta forma não consta da listagem de palavras-tema, uma

vez que, em todas as ocorrências, se constata que nem se trata do imperativo do verbo entrar nem do presente do conjuntivo do verbo entrar, uma vez que estamos perante a preposição “entre” (veja-se a figura 5).

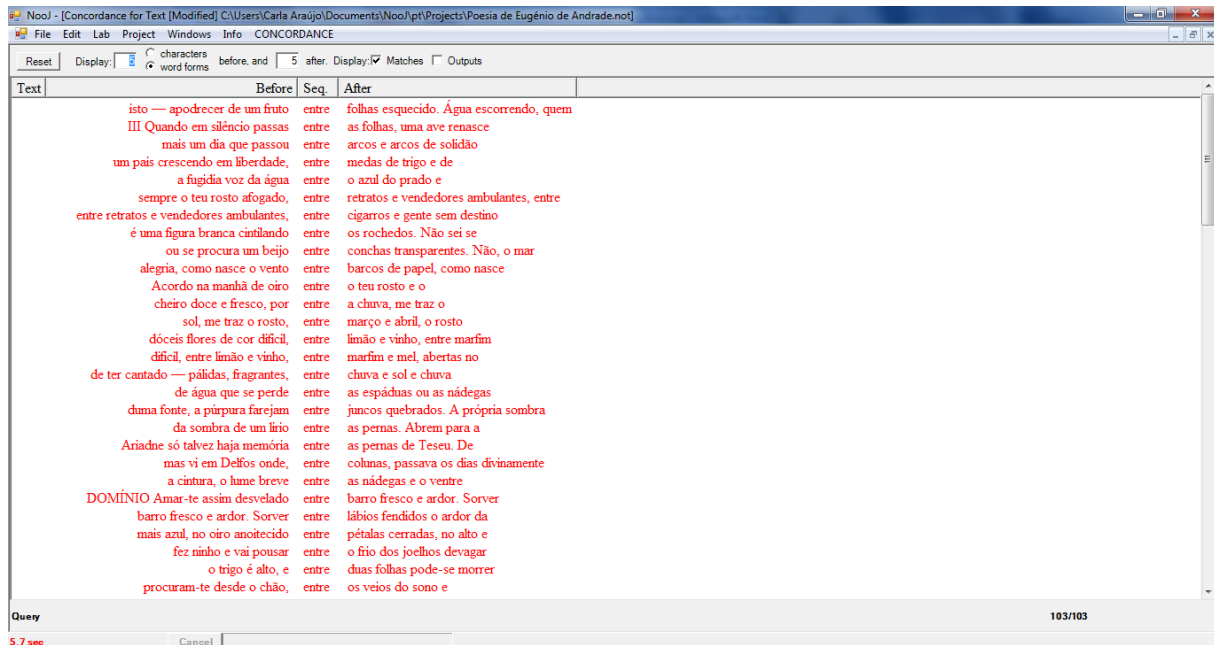


Figura 5: Concordâncias da palavra-tema “entre”.

Fonte: Nooj.

A palavra-tema “sua” configura, igualmente, um fenómeno de homonímia parcial, quer dizer, “sua” [verbo suar] e “sua” [determinante e pronome possessivo]. A análise das concordâncias permitiu-nos confirmar que a forma “sua” não é a forma do verbo suar, no presente do indicativo ou no imperativo, tratando-se do determinante e pronome possessivo, como se pode observar na figura 6.

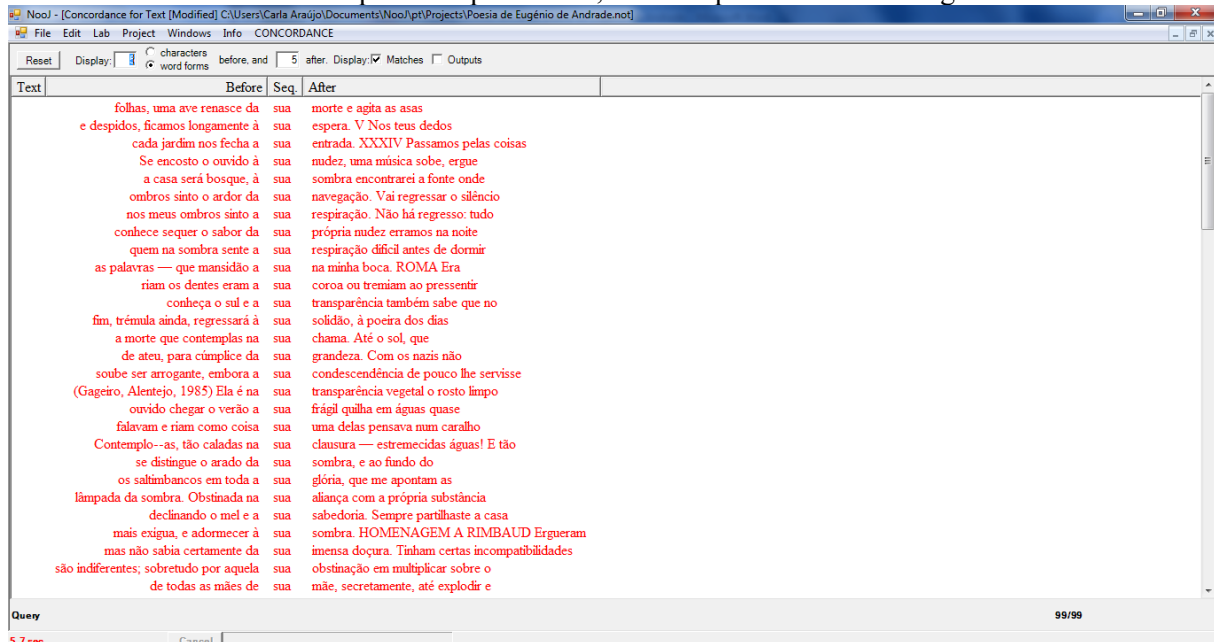


Figura 6: Concordâncias da palavra-tema “sua”.

Fonte: Nooj.

Similarmente, a palavra-tema “como” remeteu-nos para mais um fenómeno de homonímia parcial, ou seja, “como” [conjunção], “como” [advérbio interrogativo], “como” [nome masculino] e

“como” [verbo]. Como se pode observar na figura 7, em nenhuma concordância, a palavra-tema “como” corresponde à forma do verbo comer.

Text	Before	Seq.	After
	fica a vida suspensa. E	como	se um rio cantasse: em
	tremem maduras todas as espigas	como	se o próprio dia as
	-te a face. IV Somos	como	árvores só quando o desejo
	no chão um corpo aberto	como	os animais. IX. MADRIGAL Tu
	quando anoitece. Era o corpo	como	um rio em sereno desafio
	as margens quando desce. Andava	como	quem passa sem ter tempo
	os erguia no ar. Sorria	como	quem dança. E desfôlhava ao
	rítmo da própria primavera, e	como	as flores e os animais
	seu destino de cavalo indomável	como	um rio; suspensas, as aves
	Impetuoso, o teu corpo é	como	um rio onde o meu
	de folhas verdes, tão jovens	como	tu, minha alegria. Terra onde
	onde meus olhos bebem, fundo,	como	quem bebe a madrugada. XXX
	na terra de sabor amargo	como	o coração dos frutos bravos
	coisas sem as ver, gastos	como	animais envelhecidos; se alguém chama
	nos pede amor não estremece:	como	frutos de sombra sem sabor
	palavra amadureça e se desprenda	como	um fruto ao passar o
	de pedra por irmão. Tinham	como	toda a gente o milagre
	tresmalhados. Tinham fome e sede	como	os bichos, e silêncio à
	a manhã feliz e descuidada,	como	só a manhã pode brincar
	dos animais. Maio trouxe cravos	como	outrora, cravos morenos, como tu
	cravos como outrora, cravos morenos,	como	tu dizias, mas cada hora
	GAIVOTAS Contente de me dar	como	as gaivotas bebo o outono
	mais que digam é perfeito	como	a vida. Tenho tristezas como
	como a vida. Tenho tristezas	como	toda a gente. E como
	como toda a gente. E	como	toda a gente quero alegria
	retrato da moldura. Se soubesses	como	ainda amo as rosas, talvez
	o coração rosas tão brancas	como	as que tens na moldura
	coisa que vive, e é	como	um perdão que não tive

Figura 7: Concordâncias da palavra-tema “como”.
Fonte: Nooj.

Para terminar, referimos mais um caso de homonímia parcial com que nos deparamos ao analisar a lista de palavras mais frequentes, “era” [verbo ser] e “era” [nome]. De acordo com a análise das concordâncias da palavra-tema “era” (veja-se a figura 8), esta palavra-tema não integra a listagem das palavras-tema, porque se trata de uma forma do verbo ser.

Text	Before	Seq.	After
	segua o seu caminho, porque	era	um deus que passava. Alheio
	surgir. Brotou água onde tudo	era	secura. Paz onde morava a
	ELEGIA CHAMADA DOMINGO O domingo	era	uma coisa pequena. Uma coisa
	não via. O corpo todo	era	a saudade de alguém que
	pulsar do próprio coração. Moreno,	era	a forma das pedras e
	minha mãe cantar. A mãe	era	nesse tempo uma mulher triste
	ao descer. ELEGIA Às vezes	era	bom que tu viesses. Falavas
	sobre o mar cada palavra	era	apenas o rumor de um
	para dar um ao outro;	era	como se todas as coisas
	coisas eram possíveis. Mas isso	era	no tempo dos segredos, era
	era no tempo dos segredos,	era	no tempo em que o
	em que o teu corpo	era	um aquário, era no tempo
	teu corpo era um aquário,	era	no tempo em que os
	viesses e te levasse, eu	era	só fome o que sentia
	ao teu retrato como quando	era	pequeno: recortando aqui, colando ali
	onde o rumor da sombra	era	um rumor de festa, pela
	perguntava porque partia, porque ficava;	era	pouco o que tinha, pouco
	pinheiros; não dos lílases, pois	era	cedo ainda para mostrarem o
	Shakespeare MAR DE SETEMBRO Tudo	era	claro: céu, lábios, areias. O
	aman, exaltam o silêncio. Tudo	era	claro, jovem, alado. O mar
	flores de cristal. E tudo	era	água, água, desejo só de
	pedra?, pedra ou flor? Verde	era	a luz, e a espuma
	nos meus ombros. Sem muros	era	a terra, e tudo ardia
	ou uma lágrima: nada mais	era	quando despertava. 2 Sabor de água
	próximo da nascente. 4 Só água	era	, e sem memória. 5 Claridade sem
	esta carícia onde o desejo	era	só brisa? CANTE JONDO A
	viu em Delos, assim esbelto	era	o dia em que te
	que te encontrei; assim esbelta	era	a noite em que te

Figura 8: Concordâncias da palavra-tema “era”.
Fonte: Nooj.

Tendo por base a referida metodologia, como podemos observar na tabela 3, na “Poesia”, de Eugénio de Andrade, foi possível delimitar oito campos temáticos de diferentes domínios: o campo

temático do domínio de Corporalidade foi delimitado através das palavras-tema Corpo/corpo, Olhos/olhos, Boca/boca, Mãos/mãos, Mão/mão, Coração/coração, Lábios/lábios, Rosto/rosto, dedos, Olhar/olhar; o campo temático de Cores foi delimitado através das palavras-tema Azul/azul, Branco/branco; o campo temático do domínio de Temporalidade, a partir das palavras-tema Verão/verão, Tempo/tempo, Dia/dia, Dias/dias, Tarde/tarde, Manhã/manhã, Noite/noite, Vez/vez, Vezes/vezes; o campo temático de Natureza, relacionado com as palavras-tema Água/água, Luz/luz, Sol/sol, Mar/mar, Terra/terra, Vento/vento, Rio/rio, Águas/águas, Chuva/chuva, ar, Neve/neve, Folhas/folhas, Flor/flor, céu, Pedra/pedra; o campo temático do domínio de Sentimento foi construído a partir das palavras-tema Amor/amor, Alegria/alegria, Desejo/desejo, ardor; o campo temático do domínio de Declínio foi delimitado através das palavras-tema Fim/fim, Morte/morte, Morrer/morrer, Lume/lume, Silêncio/silêncio, Sombra/sombra, Só/só; o campo temático do domínio da Musicalidade foi delimitado através das palavras-tema Palavras/palavras, Palavra/palavra, Nome/Nome, Rumor/rumor, Música, Voz/voz; o campo temático de Infância foi delimitado através das palavras-tema Casa/casa, Lugar/lugar, Mãe/mãe, Ser/ser.

Tabela 3: Campos temáticos de “Poesia”, de Eugénio de Andrade.

Campos temáticos	PALAVRAS-TEMA
Corporalidade	Corpo/corpo, Olhos/olhos, Boca/boca, Mãos/mãos, Mão/mão, Coração/coração, Lábios/lábios, Rosto/rosto, dedos, Olhar/olhar
Cores	Azul/azul, Branco/branco
Temporalidade	Verão/verão, Tempo/tempo, Dia/dia, Dias/dias, Tarde/tarde, Manhã/manhã, Noite/noite, Vez/vez, Vezes/vezes
Natureza	Água/água, Luz/luz, Sol/sol, Mar/mar, Terra/terra, Vento/vento, Rio/rio, Águas/águas, Chuva/chuva, ar, Neve/neve, Folhas/folhas, Flor/flor, céu, Pedra/pedra
Sentimento	Amor/amor, Alegria/alegria, Desejo/desejo, ardor,
Declínio	Fim/fim, Morte/morte, Morrer/morrer, Lume/lume, Silêncio/silêncio, Sombra/sombra, Só/só
Musicalidade	Palavras/palavras, Palavra/palavra, Nome/Nome, Rumor/rumor, Música, Voz/voz
Infância	Casa/casa, Lugar/lugar, Mãe/mãe, Ser/ser

Fonte: Elaboração própria.

4 Conclusões

A partir da análise das palavras-tema na “Poesia”, de Eugénio de Andrade, pôde-se observar não só a sofisticação da criatividade linguística de Eugénio de Andrade e a diversidade temática da sua obra, como também a riqueza lexical da sua poesia, uma incomparável obra de arte, de extremo valor literário e, particularmente, linguístico.

Como já apresentamos na seção do estudo lexical da poesia de Eugénio de Andrade, na “Poesia”, de Eugénio de Andrade, foi possível delimitar oito campos temáticos de distintos domínios: o campo temático do domínio de “Corporalidade”, o campo temático de “Cores”, o campo temático do domínio de “Temporalidade”, o campo temático de “Natureza”, o campo temático do domínio de “Sentimento”, o campo temático do domínio de “Declínio”, o campo temático do domínio da “Musicalidade” e o campo temático de “Infância”.

Esperamos ter contribuído para complementar os estudos da poesia de Eugénio de Andrade, através de recursos computacionais, ainda pouco explorados na análise de obras literárias, mas que poderão enriquecer as análises de pendor mais tradicional, abrindo caminho para uma cooperação que

alinha a Linguística de Corpus e os Estudos Literários, e oferecendo novas alternativas para os estudos na interface entre língua e literatura.

Referências

- Andrade, E. de (2017). *Poesia*. Porto. Assírio & Alvim.
- Anthony, L. (2022). *AntConc. A freeware corpus analysis toolkit for concordancing and text analysis*. Tokyo. Japan. Waseda University.
- Anthony, L. (2023). *Estatísticas comuns usadas na linguística de corpus*. Disponível em https://laurenceanthony.net/resources/statistics/common_statistics_used_in_corpus_linguistics.pdf.
- Berber Sardinha, T. (2004). *Linguística de Corpus*. Barueri, SP. Manole.
- Berber Sardinha T. (2019). *Using multi-dimensional analysis to detect representations of national identity*. In: Berber Sardinha, T. & Veirano Pinto, M. (Eds.). *Multi-Dimensional Analysis: Research Methods and Current Issues*. London. Bloomsbury.
- Berber Sardinha, T. (2020). *Lexicogrammar*. (C.A. Chapelle, Org.). *The Concise Encyclopedia of Applied Linguistics*. Wiley Blackwell.
- Biber, D., Conrad, S. & Cortes, V. (2004). *If you look at...: Lexical bundles in university teaching and textbooks*. *Applied Linguistics*, v. 25, n.º 3, p. 371-405.
- Biber, D. (2011). «Corpus linguistics and the study of literature: Back to the future?» *Scientific Study of Literature*, 1, pp. 15–23.
- Brait, B. (2000). *Língua e Literatura: uma falsa dicotomia*. *Revista da Anpoll*, n.º 8, p. 187-206.
- Brezina, V. (2016). *Collocation networks*. In: Baker, P. & Egbert, J. (Eds.). *Triangulating methodological approaches in corpus linguistic research*. New York. London. Routledge, p. 102-119.
- Chaves, R. P. (2013). “Organização do léxico”. In: Raposo, E. (Org.) (2013). *Gramática do português*. Lisboa. Fundação Calouste Gulbenkian, pp. 185-212.
- Finatto, M. J. B., Lopes, L. & Silva, A. C. (2015). *Processamento de linguagem natural, linguística de corpus e estudos linguísticos: uma parceria bem-sucedida*. *Domínios de Linguagem*, v. 9, n. 5, p. 41–59. Disponível em: <https://lume.ufrgs.br/handle/10183/169398>.
- Firth, J. R. (1957b). A synopsis of linguistic theory, 1930–1955. *Studies in Linguistic Analysis* (pp. 1–32). Special Volume, Philological Society.
- Hoey, M., Mahlberg, M., Stubbs, M. & Teubert, W. (2007). *Text, Discourse and Corpora*. London. Continuum International Publishing Group.
- Leech, G. & Short, M. (1981). *Style in fiction*. London & New York. Longman.
- Mateus, M. H. M. (2003). *Gramática da Língua Portuguesa*. Lisboa. Caminho.
- Scott, M. (1997). *PC analysis of key words – And key key words*. *System*, v. 25, n.º 2, p. 233-245.
- Scott, M. (2022). *WordSmith Tools versão 8*. Stroud. Lexical Analysis Software.
- Semino, E. & Short, M. (2004). *Corpus Stylistics: Speech, Writing and Thought Presentation in a Corpus of English Writing*. London. Routledge.
- Silberztein, M. (2003). *Nooj Manual*. Internet. Disponível em <http://www.nooj4nlp.net/NooJ%20Manual.pdf>
- Silberztein, M. (2016). *La formalisation des langues: l’approche NooJ. Formalizing Natural Languages: the NooJ Approach*. Wiley Eds.
- Stubbs, M. (2018). *The (very) long history of copora, concordances, collocations and all that*. In: A. Cermáková, M. Mahlberg (Eds.). *The Corpus Linguistics Discourse: In honour of Wolfgang Teubert*. Amsterdam. John Benjamins.
- Zyngier, S., Viana, V., Silveira, N. G. (2011). *Discurso literário e linguística de corpus: uma visão empírica*. *Caderno de Letras*, 28, 99-107.