

O HOMEM, SEU MUNDO E SEU UNIVERSO LEXICAL. CONSIDERAÇÕES SOBRE LÉXICO E LITERATURA.

Darcilia Marindir Pinto Simões
CNPq/UERJ/FAPERJ/SELEPROT/SBPL
darcilia.simoies@pq.cnpq.br

RESUMO: Objetiva-se tratar da importância da leitura para ampliação do domínio lexical e, conseqüentemente, expansão do universo sociocultural do leitor. Por meio da exploração de textos clássicos, pretende-se não apenas desenvolver o hábito da leitura, mas, sobretudo, destacar a relevância do texto literário, em especial os contos de Eça de Queiroz, como meio de aquisição de conhecimento linguístico, histórico, político, filosófico etc.

Trata-se de relato de pesquisa em andamento, combinada com projeto de Iniciação Científica, na qual se faz um levantamento das palavras e expressões nominais presentes nos contos de Eça de Queiroz, com vistas à produção de um glossário.

Após a produção da lista de palavras e respectivas frequências, pretende-se produzir uma relação com as palavras e expressões-chave nominais que possam caracterizar a presença da ironia e que, portanto, funcionem como âncoras textuais.

Temos a crença de que a discussão desses itens léxicos, sua frequência e sua iconicidade poderão servir de base à produção de material didático que venha a estimular a leitura de outros textos clássicos a partir da constatação da influência destes não só na ampliação do repertório dos leitores, mas também no seu conhecimento de mundo.

A pesquisa vem sendo realizada com a participação de três acadêmicos¹. Partiu-se do estabelecimento do *córpus* com base na compilação das edições portuguesa de 1902 e brasileira de 1986. Em seguida, foram produzidas as listas de palavras (*wordlist*), os colocados (*collocates*) e os cotextos (*clusters*), a partir do que estamos classificando os itens considerados significativos a partir de sua frequência no *córpus*, para, em seguida, classificá-los segundo sua iconicidade em relação à produção da ironia nos contos.

Portanto, a metodologia de trabalho é a pesquisa bibliográfica combinada com o processamento digital.

Cada acadêmico-pesquisador ficou incumbido de preparar o estudo de três dos doze contos.

No momento, estão sendo concluídos os trabalhos com base digital no programa *WordSmith Tools*, pois tivemos um relativo atraso por causa da demora na entrega do volume de contos publicado em 1902.

Como se trata de pesquisa em andamento, não temos ainda resultados a apresentar.

As aulas de língua deveriam prestigiar o estudo do vocabulário, promovendo maiores oportunidades de leitura e discussão de textos e explorando as relações possíveis entre a organização textual, as prescrições gramaticais e o potencial icônico dos textos. Isto é, explorar a faculdade dos textos gerarem imagens na mente interpretadora e, por conseguinte, ativarem processos cognitivos que podem orientar/desorientar o processo de leitura e compreensão.

Tais dados deverão possibilitar a inferência das isotopias subjacentes aos contos, a partir das quais o leitor não só aprecie e passe a gostar da leitura do texto clássico, mas também amplie seu repertório verbal e enciclopédico.

Ao final da pesquisa pretende-se produzir um glossário com os itens nominais apurados nos contos estudados.

PALAVRAS-CHAVE: literatura, domínio lexical, conhecimento enciclopédico

¹ Bruna Damiana de Sá Mottinha, Bolsista CNPq; Suzana Lopes de Oliveira Gomes, Bolsista UERJ e Yuri Nikolai de Souza Teixeira, voluntário.

ABSTRACT: The objective is to demonstrate the importance of reading for the expansion of the lexical domain and hence of the socio-cultural universe of the reader. Through exploration of classic texts, it is intended not only to develop the habit of reading, but above all, emphasize the relevance of the literary text, especially the tales by Eça de Queiroz, as a means of acquisition of linguistic knowledge, historical, political, philosophical, etc.

This article reports on the ongoing research, combined with a Scientific Initiation project, in which it surveys the words and noun phrases present in the tales by Eça de Queiroz, in order to produce a glossary. After producing the list of words and their frequencies, we intend to produce a list with key words and expressions that may characterize the presence of irony and, therefore, act as anchor text.

We believe that the discussion of these tokens, their frequency and their iconicity, may serve as a basis for the production of didactic material that will stimulate others to read classical texts from the observation of their influence not only in expanding the repertoire of readers, but also on their knowledge of the world.

The survey is being conducted with the participation of three scholars [1]. The starting point was the establishment of the corpus based on the compilation of Brazilian (1986) and Portuguese (1902) issues. They produced some lists of words (wordlist), the put (collocate) and the co-text (clusters), from what we are sorting the items considered significant from their frequency in the corpus, to, then, classify them according to their iconicity in relation to the production of irony in the tales. Therefore, the methodology of work is the literature combined with digital processing.

Each scholar-researcher was commissioned to prepare the study of three of the twelve tales.

Currently, the work is being completed based digital program WordSmith Tools, for we have had a late because of delay in delivery of the tales collection, published in 1902.

As this is ongoing research, we don't have yet how produce results.

The language lessons should honor the study of vocabulary, promoting greater opportunities for reading and discussion of texts and explore possible relationships among the textual organization, grammar standard and potential iconic of texts. This is: exploring the possibility of generating images in interpreting mind and therefore activate the cognitive processes that can orient/disorient the process of reading and understanding.

Such data should lead to the inference of the isotopies behind the stories, from which the reader not only enjoy, but enjoy reading the classic texts and also broaden its verbal and encyclopedic repertoire.

At the end of the research aims to produce a glossary of the nominal items verified in short stories studied.

KEYWORDS: literature, lexical field, encyclopedic knowledge.

Apresentação

Dividimos esta comunicação em três etapas. A primeira é a exposição da fundamentação teórica do projeto de produtividade em pesquisa denominado “Iconicidade em Eça de Queiroz: leitura para o domínio da língua”, iniciado em 2010 e com término previsto para 2014. A segunda etapa é a descrição do projeto em si, seus objetivos e produtos pretendidos. Por fim, apresentar uma breve demonstração do trabalho em realização.

Base teórica

Subsidiomos a pesquisa da qual trazemos breve relato com a Teoria da Iconicidade Verbal — TIV— (SIMÕES, 2007, 2009, 2010) em diálogo com os Gêneros do Discurso — GD — (BAKHTIN, 1997) e a Gramática Sistêmico-funcional de Halliday — GST— (2004). Também buscamos apoio na ferramenta digital

WordSmith Tools 5.0, para automação do levantamento do vocabulário e de seus contextos imediatos (cotextos).

A TIV funda-se na Semiótica de Peirce e busca descrever o signo e sua movimentação no texto de modo a percebê-lo como elemento orientador ou desorientador da compreensão. Tem-se na TIV, na perspectiva diagramática sintagmática, em especial, uma significativa contribuição para o entendimento do papel dos signos no interior dos textos. A partir da relação entre o signo e seu objeto imediato (aquele ao qual se refere), pode-se estabelecer uma relação icônica, indicial ou simbólica. Destas é possível deduzir vários níveis de entendimento do texto: a) Gênero predominante; b) Variedade linguística; c) Recortes temáticos; d) Coesão; e) Coerência; f) Clareza; g) Adequação lexical etc.

Para tanto, a TIV aponta a existência de itens léxicos que funcionam como *palavras-chave* — itens que garantem a manutenção temática — e *âncoras textuais* — palavras-chave que funcionam como nós sêmicos e servem para definir a(s) isotopia(s) subjacente(s) ao texto. Quando se trata de texto literário, que é polissêmico, podem ser percebidas várias isotopias entrecruzando-se na superfície textual. Nesse caso, as âncoras textuais são as palavras-chave que participam simultaneamente de mais de uma isotopia, representando assim os pontos em que a trama temática se instala e é garantida a existência do texto. Ao contrário, nos textos referenciais (técnicos e científicos), espera-se uma única isotopia, portanto, as âncoras textuais (se as houver) cruzar-se-ão sobre si mesmas, exclusivamente para amarrar o texto.

Do léxico e do conhecimento de mundo

Trazemos palavras de Tania Câmara, para reforçar o que pensamos sobre léxico.

Numa perspectiva cognitivo-representativa, o léxico constitui a codificação da realidade extralinguística interiorizada de uma comunidade linguística. Quanto maior for o domínio lexical do usuário, maior será sua capacidade de interação com o outro. (In *Revista Philologus*, Ano 16, N° 46. jan./abr.2010)

Somando a essa idéia a ótica sistêmico-funcional, infere-se que as metafunções ideacional e interpessoal atingirão graus de maior eficiência tanto maior seja o repertório do falante. A função textual dispensa comentários, uma vez que é obviamente construída com palavras. Veja-se que repertório pode ser definido como o conjunto de vocábulos de uma língua ou como o domínio vocabular de um sujeito. Em um e outro casos, verifica-se a relevância da competência lexical como sendo base para uma expressão eficiente.

Assim sendo, cumpre que a prática pedagógica de uma língua promova o exercício da leitura como fonte de aquisição de itens léxicos. Ainda que a leitura dos textos clássicos não mais seja o foco das aulas de língua portuguesa no Brasil, cumpre lembrar que tais textos são verdadeiro manancial de informações lexicais, sem contar com o volume de dados enciclopédicos ali existentes. Se hoje se constata certa indigência verbal nos alunos do Ensino Médio e da graduação, pode-se atribuir esse problema à falta de leitura em geral e da literatura clássica, em especial.

Dizer-se que o estudante não gosta de ler por força de um modelo de sociedade que prioriza a leitura, não mais é suficiente, pois a sociedade digital vem trazer mil e um recursos que de alguma forma facilitam o acesso ao texto literário e, se provocado, o estudante entrará em contato com essa literatura e, se bem orientado, descobrirá nesse material um universo fascinante, como comprovam os relatórios dos

acadêmicos² de Iniciação Científica que participam do projeto em questão. Segundo eles, aproximar-se da obra de Eça de Queiroz e realizar as leituras correlatas tem resultado imediato na melhoria de seu desempenho nas avaliações das disciplinas em curso. Logo, o jovem estudante gostará de ler o clássico quando este lhe for apresentado de modo adequado e oportuno.

Um dos grandes ganhos com a leitura dos contos-cópus de nossa pesquisa é o conhecimento de outra época, outros usos e costumes e tudo isso materializado em itens léxicos novos para os leitores. Portanto, cada página lida se mostra como uma aventura, e o mundo do leitor vai-se alargando na proporção em que a leitura avança.

A pesquisa lexical nos contos ecianos

Depois de estabelecido o texto por meio de cotejo entre as edições³ portuguesa de 1902 e brasileira de 1986, processaram-se os contos com o programa digital WordSmith Tools 4.0, com o qual foram produzidas as listas de palavra presentes em cada conto e, posteriormente, uma grande lista com todas as palavras atualizadas nos contos. Construímos uma trilha pelo léxico para estimular a busca do clássico literário.

Iniciamos a valorização do olhar para o texto literário, trazendo à cena palavras de Roland Barthes:

Entendo por *literatura* não um corpo ou uma seqüência de obras, nem mesmo um setor de comércio ou de ensino, mas o grafo complexo das pegadas de uma prática: a prática de escrever. Nela viso, portanto, essencialmente, o texto, isto é, o tecido dos significantes que constitui a obra, porque o texto é o próprio aflorar da língua, e porque é no interior da língua que a língua deve ser combatida, desviada: não pela mensagem de que ela é o instrumento, mas pelo jogo das palavras de que ela é o teatro. Posso, portanto dizer, indiferentemente: literatura, escritura ou texto. (Aula, 1977)

Partindo dessa perspectiva, trouxemos os contos de Eça para a pesquisa, com o fito de não apenas desenvolver mais um estudo da língua portuguesa numa ótica diatópica e diacrônica a um só tempo, mas, sobretudo, dirigir o olhar para o fazer literário como a arte da palavra a reinauguração do código. Então, buscamos nas listas produzidas com a ajuda digital itens léxicos que pudessem comprovar a importância da escolha para a tessitura do texto. Destacar a seleção lexical como signo da identidade do autor, uma vez que o desenho de seu estilo se constrói com e nas palavras.

Uma vez em contato com as técnicas de identificação e análise dos itens léxicos que funcionam como âncoras textuais e constituem trilhas para a leitura, o trabalho com os clássicos literários deixa de corresponder a uma experiência negativa, passando a ser visto como desafiador. Ao aplicar a *Teoria da Iconicidade Verbal* na leitura dos textos literários constrói-se uma base para melhor entendimento do texto e interpretação mais acurada das nuances encontradas na materialidade textual.

A tarefa de produção das listas realizou-se em dois momentos: (1) produção automática de doze listas de palavras, referentes a cada um dos doze contos; (2) a elaboração de uma lista única a partir do

² Duas bolsistas e um voluntário.

³ QUEIROZ, Eça de. *Eça de Queiroz. Obra completa*. Rio de Janeiro: Nova Aguilar, 1986.

-----, *Contos. Prosas Bárbaras*. Porto: Lello & Irmãos Editora, 1902.

processamento de todos os contos em um mesmo arquivo. Os objetivos imediatos de ambos eram semelhantes: (a) observar quais são os substantivos e os adjetivos de maior ocorrência em cada conto; (b) observar quais são os substantivos e os adjetivos de maior ocorrência na totalidade dos contos; (c) analisar se os conjuntos de nomes de maior ocorrência contemplam os temas, subtemas e isotopias encontrados nos textos. A fim de contemplar a média de 85% do vocabulário total de cada conto, foi estabelecido o corte mínimo em cinco ocorrências.

Vejam-se a seguir exemplos três das doze tabelas produzidas com as palavras de cada conto, processado individualmente. Os dados das tabelas são:

- coluna 1 representa o número de ordem em que a palavra aparece na lista geral processada;
- coluna 2 traz a relação de palavras em foco
- coluna 3 contém o número de ocorrências:

Vamos às tabelas.

Singularidades de uma Rapariga Loira			Civilização			Adão e Eva no Paraíso		
Nº	Substantivo	Ocorrências	Nº	Substantivo	Ocorrências	Nº	Substantivo	Ocorrências
09	MACÁRIO	112	14	JACINTO	75	17	ADÃO	68
32	LUÍSA	35	37	VIDA	23	28	PAI	43
34	TIO	34	43	TORGES	19	36	EVA	33
40	FRANCISCO	23	47	SÉCULO	17	38	PARAÍSO	31
42	CASA	20	50	AMIGO	15	44	TERRA	27
47	DIA	19	53	BRÁS	14	51	HOMEM	23
48	SENHOR	19	54	CIVILIZAÇÃO	14	60	PAIS	19
49	AMIGO	18	61	ÁGUAS	12	64	OLHOS	18
52	CAIXEIRO	16	65	HOMEM	12	75	RIO	13
53	HOMEM	16	66	SR	12	77	ÁGUA	12
56	JANELA	15	70	ZÉ	12	78	ÁGUAS	12
58	MÃE	14	73	MESA	11	80	ÁRVORE	12
59	MÃO	14	80	FONÓGRAFO	10	81	CAVERNA	12
62	NOITE	13	82	MÃOS	10	83	PATAS	12
67	PEÇA	12	86	SOL	10	86	SOL	12
70	TEMPO	12	87	VINHO	10	87	TERROR	12
72	VIDA	12				91	CARNE	11
73	VOZ	12				93	LUME	11
83	MULHER	10				94	MAR	11
86	QUARTO	10				97	SANGUE	11
87	SENHORA	10				101	FLORESTA	10
						102	GALHOS	10

Ainda estamos no estágio de exploração da ferramenta *WordList*. Já estão feitas as listas de palavras individuais (*wordlist*) e a lista de múlti-palavras (*wordlist, clusters activated*). Falta-nos fazer as listas de palavras de consistência individuais (*detailed consistency*), a lista de múlti-palavras2 de consistência (*detailed consistency, clusters activated*) e, por fim, faremos a lista de dimensões e densidade lexical (*statistics*). Esta nos ajudará a classificar, principalmente, as âncoras textuais.

Posteriormente, com a ferramenta *Concord*, produziremos novas tabelas apresentando: 1- Concordância (*concordance*); 2 - lista de colocados (*collocates*) e 3 - lista de agrupamentos lexicais⁴ (*clusters*) e 4 - lista de padrões de colocados (*patterns*).

⁴ Cf. *lexias complexas* no Hyperbase.

Optamos pela ferramenta de processamento digital com vista a conferir um carácter menos subjetivo à discussão dos itens vocabulares, uma vez que estaremos nos apoiando em quantitativos e estatística. Essa opção também decorre do desejo de levar o jovem pesquisador a desfazer a ideia de que em língua tudo é subjetivo e cada um diz uma coisa. Por meio das ferramentas digitais, dá-se ao trabalho uma feição matemática que é mais convincente, em especial, ao estudante jovem.

Por que discutir vocabulário em texto artístico?

Olhado na perspectiva da produção, temos que, no texto literário, o signo verbal, que é convencional por origem, ganha funções icônicas e indiciais segundo a habilidade do autor. Segundo a ótica do intérprete (ou leitor, em outras bases teóricas), este se tornará mais competente, pois a orientação sobre as especialidades sígnicas (icônicas, indiciais e simbólicas), possibilita uma análise em que o componente lexical seja observado não apenas com as lentes do dicionário, mas também com as lentes oferecidas pelo próprio texto e pelo contexto. Desta forma, entendemos que o falante além de enriquecer seu vocabulário (saber linguístico) evolui como observador e como comunicador ao lidar com textos, em geral; e o vocabulário e as informações enciclopédicas oriundas dos textos clássicos podem promover a tão almejada interdisciplinaridade, quando se faz um trabalho articulado entre língua portuguesa e outras disciplinas.

O estudo dos signos em perspectiva semiótico-funcional não desautoriza a gramática, senão ao contrário comprova sua importância e permite seja entendida a estruturação gramatical como o suporte invisível que sustenta o texto. A construção desse suporte será tão mais bem sucedida quanto mais bem feita seja a seleção dos itens léxicos atualizados no texto. Portanto, o trabalho do léxico na perspectiva da teoria iconicidade verbal se torna ferramenta indispensável na aquisição de domínio da língua em sua amplitude.

Para ilustrar, temos:

D. MARIA DA PIEDADE era considerada em toda a vila como “uma senhora modelo”. O velho Nunes, diretor do correio, sempre que se falava nela, dizia, acariciando com autoridade os quatro pêlos da calva:

– A vila tinha quase orgulho na sua beleza delicada e tocante; era uma loura, de perfil fino, a pele ebúrnea, e os olhos escuros de um tom de violeta, a que as pestanas longas escureciam mais o brilho sombrio e doce. (Eça de Queiroz, “No Moinho”)

Tomando-se o item *vila* como referência de análise lexical, verifica-se sua participação na apresentação da personagem Maria da Piedade: 1) era uma senhora modelo *entre aspas*; 2) era motivo de *quase orgulho*. De posse desses mínimos dados já se prenuncia uma personagem que será tratada mediante o conflito entre o ser e o parecer.

Frei Genebro, na verdade, completara a perfeição em todas as virtudes evangélicas. Pela abundância e perpetuidade da Oração, ele arrancava da sua alma as raízes mais miúdas do Pecado, e tornava-a limpa e cândida como um desses celestes jardins em que o solo anda regado pelo Senhor, e onde só podem brotar açucenas. A sua penitência, durante vinte anos de claustro, fora tão dura e alta que já não temia o Tentador; e agora, só com o sacudir a manga do hábito, rechaçava as tentações, as mais pavorosas ou as mais deliciosas, como se fossem apenas moscas importunas.

(Eça de Queiroz, “Frei Genebro”)

Observando a personagem de um Frei que “completara a perfeição em todas as virtudes evangélicas”, vê-se no emprego de itens léxicos como *abundância, manga do hábito, moscas*, por exemplo, verifica-se uma descrição irônica da conduta do religioso.

Veja-se a tabela de nomes substantivos e adjetivos com a ocorrência mínima de 5 empregos:

18	GENEBRO	24	Observe-se que comparado o número de ocorrências de Genebro, Senhor e Deus, veremos que o frei é maior que Senhor e, conseqüentemente, maior que Deus, pois as ocorrências consecutivas são: 24, 16 e 11.
21	FREI	18	
22	ELE	18	
23	IRMÃO	17	Continuando a observação dos quantitativos, verifica-se que Genebro, Frei e Ele somam 60 ocorrências, contra a soma de 27 para Senhor e Deus.
24	EGÍDIO	17	
25	SENHOR	16	
30	ALMA	14	Ora, se a seleção lexical é uma das formas de materialização verbal do projeto de dizer do enunciador, pode-se inferir que nesse projeto a personagem Frei Genebro (que dá nome ao conto) é, de fato, o foco da atenção e, a despeito de pretender-se apresentá-lo em sua perfeição, as ocorrências levantadas já denunciam a superioridade com que se apresenta o Frei em relação ao Senhor Deus.
31	BOM	13	
36	DEUS	11	
37	MÃO	10	E por aí vai.
40	PRATO	10	
41	ANJO	9	
50	CORPO	7	
52	ERMITÃO	7	
57	PEDAÇO	6	
58	POBRE	6	
59	PORCO	6	
62	SOMBRA	6	
65	AMOR	6	
66	ANJOS	6	
68	BOAS	6	
69	BRAÇOS	6	
71	CARIDADE	5	
72	COLINA	5	
74	MARIA	5	
76	SANTA	5	
78	TERRA	5	

Finalizando, pensamos estar construindo um material técnico-teórico que venha a subsidiar as aulas de compreensão e interpretação de textos de modo a instrumentalizar os estudantes para um trabalho consciente com os signos, de modo a, quando instados a produzir textos possam lançar mão de estratégias observadas em textos como os de Eça que são, indiscutivelmente, astuciosamente construídos.

Referências bibliográficas.

- BAKHTIN, M. *Estética da criação verbal*. Os gêneros do discurso. 2ª ed. São Paulo: Martins Fontes, 1997.
- BARTHES, Roland. *Aula*. Tradução e posfácio de Leyla Perrone-Moisés. São Paulo. Cultrix. s/d.
- CÂMARA, Tania Maria Nunes de Lima. “Léxico e conhecimento de mundo”. *Revista Philologus*, Ano 16, Nº 46. Rio de Janeiro: CiFEFiL, 54 jan./abr.2010
- HALLIDAY, Michael. A. K. *An introduction to functional grammar*. London: Edward Arnold. [1985] 2004.

SIMÕES, Darcilia. “Palavras: fios do texto”. In *Acta Semiotica et Linguistica.* , v.13, São Paulo: SBPL. 2010. [p.79 – 92]

_____. Iconicidade lexical. Teoria e Prática. Edição online. Rio de Janeiro: Dialogarts. 2009. In www.dialogarts.uerj.br

_____. Iconicidade e Verossimilhança. Semiótica Aplicada ao Texto Verbal. Edição online. Rio de Janeiro: Dialogarts. 2007. In www.dialogarts.uerj.br

SARDINHA, Tony Berber. “Usando WordSmith Tools na investigação da linguagem”. <http://sites.uol.com.br/tony4/homepage.html> . *DIRECT Papers 40*. ISSN 1413-442x. 1999