

ONEIROCRITICA, LIVRO 5: APLICAÇÃO DE ANÁLISE LEXICAL AUTOMÁTICA EM UM ESTUDO DE LITERATURA GREGA ANTIGA

Anise A. G. D'Orange Ferreira*

RESUMO: *Este artigo descreve a aplicação da ferramenta de análise lexical automática – WordSmith Tools – em um estudo da Oneirocritica, de Artemidoro Daldiano, que aborda o significado dos presságios no logos onirocrítico. O corpus de estudo é o quinto e último livro da Oneirocritica e o de referência compreende os cinco livros. Com as palavras-chave obtidas pela análise de frequência lexical – os itens lexicais morreu e tornou-se – levantou-se a hipótese de a morte ser um dos principais prognósticos e preocupações onirocríticos. O levantamento temático realizado com a ferramenta é demonstrado e suas possibilidades, discutidas.*

PALAVRAS-CHAVE: *Oneirocritica, Artemidoro, onirocritica, frequência lexical, WordSmith Tools, Thesaurus Linguae Graecae, TLG Workplace, leitura automática de corpus, literatura grega antiga.*

Os textos em formato digital dos corpora eletrônicos e as ferramentas de análise lexical têm tornado mais acessível uma abordagem empírica e quantitativa da língua, própria da

* Pontifícia Universidade Católica de São Paulo – PUC-SP

Linguística do Corpus,¹ baseada em amostras de diferentes registros de linguagem em uso, sejam originalmente textos, sejam transcrições de fala. Essa abordagem tem contribuído para estudos léxico-gramaticais, gramaticais, semânticos, pragmáticos, estilométricos, análises do discurso, e possui aplicações não só no ensino de língua e literatura, mas em áreas afins, como na sociolinguística, psicolinguística, psicologia social, entre outras. (McEnery & Wilson, 1996; Biber, Conrad & Reppen, 1998).

Uma aplicação tradicional dos estudos quantitativos na literatura é a da comprovação de autoria. Oakes (1998) fornece alguns exemplos relacionados à literatura grega antiga. Em um deles, a porcentagem da partícula *kai* foi adotada como critério estilístico para a determinação da autoria das epístolas de São Paulo; em outro, a diferença na frequência de palavras comuns pôde indicar se textos aristotélicos sobre ética poderiam ser parte da *Ética a Nicômaco* ou da *Ética a Eudêmo* e, ainda em outro, a ocorrência de determinadas palavras pôde sugerir a cronologia das obras de Sócrates. A cronologia obtida por meio de técnicas estilométricas tem sido alvo de atenção, como o estudo polêmico de Ledger sobre as obras de Platão (Ledger, 1989; Keyser, 1991, 1992; Ledger, 1992; Clayman, Crane & Guthrie, 1992).

A motivação para uma abordagem empírica a exemplo da Linguística do Corpus no estudo da *Oneirocritica*², obra de Artemidoro de Daldis (séc. II d.C.) sobre a interpretação de sonhos, é decorrente de duas facilidades técnico-metodológicas. Uma delas é a existência do corpus eletrônico da língua grega, o *Thesaurus Linguae Graecae* (TLG)³, que contém quase 11 mil obras e 76 milhões de

¹ Abordagem que pressupõe um corpus legível por máquina, de tamanho finito, que se constitui uma amostragem representativa de registros autênticos da linguagem em uso.

² Transliteração do nome grego da obra, que significa *Análises de sonho*. *Onir* é abreviação de *Onirocriticon*, a forma latina do título da edição crítica de R. Pack, *Artemidori Daldiani*, 1963. Os números que seguem o título referem-se ao livro e capítulo respectivamente.

³ A sigla TLG será empregada em lugar de *Thesaurus Linguae Graecae*.

palavras; a outra é a existência de ferramentas de análise lexical, como o *WordSmith Tools (WST)*⁴ (Scott, 1997). Alguém poderia dizer que a *Oneirocritica* não representa um corpus de estudo grande o suficiente para merecer uma leitura automatizada “por meio da máquina”. Por outro lado, quase 65 mil itens lexicais para um corpus parece um número suficientemente grande para justificar o auxílio de recursos mecânicos para análises mais refinadas, baseadas em cálculos de freqüências. Vale lembrar que o próprio Artemidoro se baseou no argumento da freqüência para avaliar a relevância de um sonho ao interpretá-lo.⁵

Sistemas automatizados podem ajudar a localizar indícios e evidências para esclarecer uma determinada questão científica, mas a escolha do problema, as decisões do procedimento e as análises e interpretações extraídas dos dados estão à mercê dos *insights* que o pesquisador pode obter de acordo com sua experiência e interlocução acadêmicas.

A obra de Artemidoro vem sendo estudada não só na esfera da linguagem propriamente dita, mas também na esfera das ciências sociais e do comportamento. O presente trabalho avança nos domínios da lingüística moderna e da literatura clássica, é de natureza temática, mas pode ter uma implicação histórica, uma vez suscitado a partir da posição expressa por Bowersock (1994: 98) de que:

It would be wrong to assume that the obsession with purely predictive interpretation of dreams, as we find in Artemidorus,

⁴ A sigla WST será empregada em lugar de *WordSmith Tools*.

⁵ *Onir.* 4.27 “Os sonhos vistos com freqüência, [τοὺς πολλάκις ὀραμένους ὄνειρους](...) julga[-os] significar sempre a mesma coisa; por causa disso, de serem vistos freqüentemente, que mais se deve prestar atenção e confiar neles; pois também nós, como de costume, quando temos alguma coisa importante a dizer, a dizemos com freqüência”. Tradução própria a partir do grego. Cf. as traduções nas línguas inglesa, Artemidorus, 1990; francesa, Artémidore, 1975; espanhola, Artemidoro, 1989 e italiana, Artemidoro, 1975.

represents anything commonplace or fundamental at all in the ancient world. It is a grossly illegitimate move to appropriate Artemidorus as a representative of the masses of Greeks and Romans and to impute to them protocols of sexuality or morality on the basis of his dream book. Artemidorus does indeed deserve a place in the history of the period, but his place remains as scholar of the occult, a chronicler of the unconscious mind, and a professional profeteer in the business of making forecasts.

Os sonhos premonitórios (óneiro)⁶ são rejeitados como fonte de realidade por Bowersock (1994: 83): “em lugar algum do mundo antigo”, diz ele, “pode-se encontrar alguém que ‘sonhou estar alimentando o pênis como se este fosse um animal’ (*Onir.5.62*)”. Outro indício de falta de representatividade para o historiador seria o fato de que Artemidoro almejava ter a elite como cliente. Critica, então, abertamente Winkler (1990: 11), que se refere aos clientes de Artemidoro como “average people” e culpa Foucault (1984) por ser o líder dessa linha de análise. Acrescenta que a previsão do futuro não despertava o interesse dos romancistas gregos, que estariam, de fato, preocupados com os medos e desejos dos sonhadores. Deixou implícito, desse modo, que os medos e desejos não teriam valor similar para Artemidoro em sua *Oneirocritica*, julgando que os sonhos presentes no gênero do romance greco-romano poderiam fornecer mais interesse como projeção de anseios na esfera social, principalmente em relação ao comportamento amoroso.

Sem tentar resolver a questão da representatividade da expressão de desejos e temores em relação a gêneros literários distintos, supôs-se que não era, na onirocricia de Artemidoro, o conteúdo onírico em si que projetaria ansiedades e desejos, mas toda a análise do sonho, mais precisamente o prognóstico ou, como se encontra mais sistematicamente no quinto livro da *Oneirocritica*, o relato do evento que sucedeu ao sonho, tratado como um dado

⁶ O acento agudo é empregado no presente texto para indicação de acento tônico, e o acento circunflexo, para indicação de vogal longa (ê, para êta, η, e ô para ômega, ω).

empírico. Tampouco a clientela de Artemidoro parecia pertencer a uma elite tão seleta quanto Bowersock defendia, embora as interpretações pudessem refletir os valores e preconceitos do autor em relação ao status social das pessoas de um modo geral.

O objetivo da aplicação de uma ferramenta de análise lexical foi fazer um levantamento preliminar do léxico de conteúdo que pudesse indicar marcas de medos e desejos dos sonhadores, como temas para uma análise posterior do ponto de vista literário. Era preciso uma ferramenta que fornecesse, no mínimo, a lista de todas as palavras com a respectiva frequência. Mas era desejável dispor de uma que gerasse, ainda, uma lista de palavras com frequência significativa em relação à obra de Artemidoro. Isso seria possível pelas *keywords* ou palavras-chave obtidas com o programa WST, descrito por Oakes (1998) como uma ferramenta de concordância, dotada de estatística avançada, que ajuda a caracterizar o tema ou gênero do texto. Esse programa tornou-se uma referência para a pesquisa da linguagem mediada por computador (Berber Sardinha, 1999a) e tem sido eficaz na demarcação de gênero e temas segundo a teoria da gramática sistêmico-funcional, por meio do estudo de frequências lexicais de textos na língua inglesa e portuguesa, cuja visão probabilística do uso da linguagem justifica buscar pelas escolhas típicas (Ramos, 1997; Collins & Scott, 1997; Freitas, 1997). Uma visão similar também tem motivado a criação de outras ferramentas automáticas de análise de palavras, criadas especialmente para a língua grega antiga, tais como *parsers* (Crane, 1991) e ferramentas de colocação (Rydberg-Cox, 2000), dentro corpus do *Perseus Project* (Crane, 1997), cujo acervo de textos em grego é inferior ao do *TLG* e não contém ainda a obra de Artemidoro. A lista de *keywords* no WST é criada a partir da comparação de duas listas de palavras: uma proveniente do corpus de estudo e outra proveniente de um corpus de referência, indicando, dentro do corpus de estudo, as palavras-chave que são mais frequentes do que o esperado estatisticamente em relação à frequência do corpus maior ou de referência. Como a *Onirocritica* representa a única obra de oniro-

crícia da literatura grega que sobreviveu até nossos dias, o universo considerado para o corpus de referência foram os cinco livros que compõem a obra, supondo-se haver diferenças entre livros quanto a escolhas lexicais. O quinto e último livro da *Oneirocritica* foi selecionado como corpus de estudo por ter sido esse livro designado pelo próprio autor como uma coleção de sonhos e de exemplos de interpretação, de fácil consulta e extraídos da prática.

1. Corpora

Os livros da *Oneirocritica* em formato digital foram obtidos do *TLG* inicialmente em duas versões diferentes, ambas em formato texto, em ASCII: uma obtida diretamente do sistema *Ibycus* e outra contida na versão "E" do CD-ROM. Ambas foram fornecidas pelo laboratório do Projeto *TLG*, da Universidade da Califórnia em Irvine. Esse projeto, fundado em 1972, dispõe de uma coleção digital de textos que abrangem desde Homero até 600 d.C. e alguns textos historiográficos, lexicográficos e escoliásticos do período que vai de 600 a 1453 d. C. O texto em formato digital da *Oneirocritica* corresponde à edição de Roger Pack (1963). Apesar de John Houlihan (1997) ter sugerido alterações para a uma nova edição do texto do primeiro livro a partir do estudo da tradução árabe⁷ da *Oneirocritica*, publicada após a edição de R. Pack, o corpus no presente estudo não foi alterado em função de suas propostas.

2. Ferramentas de leitura e de análise de frequências

Os textos gregos do *TLG* em formato digital requerem software para sua leitura. O corpus é composto de arquivos em formato de

⁷ Artémidore d'Éphèse, Fahd (Ed.), 1964.

texto, separados por autores, identificados pelo número de classificação no Cânon (Berkowitz & Squitier, 1990). As obras de um mesmo autor encontram-se dentro de um mesmo arquivo, marcadas por códigos ou etiquetas (*tags*) que indicam o título da obra. Outras etiquetas indicam livros, partes, páginas, parágrafos, linhas e anotações de acordo com a edição crítica correspondente, editada em papel. Os arquivos em formato de texto, com os códigos aparentes, a rigor, podem ser lidos em qualquer processador de texto, que, todavia, não podem interpretar as etiquetas empregadas no corpus do *TLG*. Com fontes latinas, lê-se o texto pelo Código Beta. Esse código é uma transliteração de caracteres específicos da língua grega, seus acentos e marcas, em caracteres ASCII. Por exemplo, a letra *a* seguida da barra, '*a/*', indica alfa com acento agudo (ᾰ); '*a*', indica alfa com espírito fraco (ᾱ); '*a=*', indica alfa com acento circunflexo (ᾶ); '*a*', indica alfa com acento grave (ᾷ); '*a|/*' indica alfa com iota subscrito e acento agudo (ᾶ̣), e assim por diante.⁸ Com fontes gregas instaladas no sistema, lê-se o texto com caracteres gregos.

A leitura dos textos, com a interpretação correta dos códigos ou etiquetas, requer um software especial.⁹ Aqui, foi adotado o *TLG Workplace*, (1997, v. 6.01; 1999, v. 7.01 e 2000, v. 8.0, para PC/Win95).¹⁰ Esse software possui funções de busca por autor, obra, período, cânon, numeração de capítulos, e realiza concordância sobre todo o corpus ou sobre obras previamente selecionadas de diferentes autores. As fontes gregas disponíveis no *TLG Workplace* para leitura de código beta do CD e conversão para caracteres gregos são *Sgreek*, *Wingreek* e *Silver Humana-Unicode*. Mas a fonte que faz a conversão direta do código beta e que pôde ser utilizada no programa de análise lexical foi *Sgreek*.

⁸ Detalhes do código podem ser obtidos em <http://www.TLG.uci.edu/BetaCode.html>.

⁹ É possível extrair os arquivos do CD pelo código do autor e lê-los em um processador de texto qualquer. A busca de obras, partes, páginas e linhas específicas só é possível por meio de um software de leitura do *TLG*.

¹⁰ Com o *TLG Workplace*, é possível recuperar uma obra, fazer busca de um item lexical por autores, data e classificação de gênero, mas não é possível extrair e salvar subcorpora compostos de obras extraídas de autores diferentes.

Para análise lexical, foi adotado o programa *WST v.2.0* (Scott, 1997), cuja descrição mais detalhada pode ser encontrada em Berber Sardinha (1999b). A ferramenta *Wordlist* contida no programa forneceu as listas de freqüência dos corpora. A ferramenta *Keywords* forneceu os dados principais, uma lista de palavras-chave, i.e., de palavras de freqüência incomum em um corpus (de estudo) quando este é comparado a um corpus maior (corpus de referência). A ferramenta *Concord* e o mecanismo de busca do *TLG Workplace* forneceram as concordâncias, ou a indicação das palavras-chave em contexto; a lista de colocados (*collocates*) e de padrões de colocados (instrumentos da ferramenta *Concord*) forneceram as co-ocorrências e posições das palavras-chave.

Keywords. No *WST*, palavra-chave (*keyword*) é definida como a palavra cuja freqüência se situa acima da freqüência esperada, tendo-se como referência uma freqüência mínima estabelecida pelo pesquisador. A freqüência mínima estipulada foi 3 e a probabilidade máxima de erro (valor α), associada ao valor crítico encontrado pelo teste estatístico, foi 0,000001. Nas ciências sociais, a probabilidade de erro aceitável ou o fator de risco calculado costuma variar de 0,01 a 0,05. Como em análise lexical, a seletividade é mais importante que o risco, o autor do software sugere um valor de probabilidade inferior, de modo a se obter um número menor de palavras-chave. O teste estatístico embutido no programa e empregado neste estudo foi o *Log Likelihood*, proposto por Ted Dunning (1993).¹¹ Ele gera automaticamente a probabilidade de ocorrência dos itens lexicais, fornecendo sua *keyness* ou chavicidade, que pode ser positiva ou negativa. Positiva quando a freqüência é maior do que a esperada em comparação com o corpus de referência (corpus maior) e negativa quando a freqüência é menor do que a esperada em

¹¹ Esse teste não pressupõe uma distribuição normal dos dados para cálculo de significância e parece igualmente eficiente para amostras pequenas. "The advantage of the G-square or log likelihood measure is that it does not assume the normal distribution" (Oakes, 1998:189).

comparação com o corpus de referência. As palavras então listadas como resultado de chavicidade são as “incomumente freqüentes”¹² em comparação com as que seriam esperadas com base no corpus maior. As palavras de freqüência baixa incomum (as *keywords* negativas) são listadas no final em cor diferenciada. Neste estudo foram analisadas algumas palavras-chave com chavicidade positiva.

3. Procedimentos

Preparação do texto digital. Alguns cuidados tiveram de ser tomados para a utilização do texto grego no *WST*. Foi necessário fazer uma padronização do texto grego nos dois corpora. Embora não houvesse nenhum problema em relação ao idioma, uma vez que o texto grego do *TLG* é uma transcrição em ASCII do Código Beta, as marcas de parágrafos, de mudança de linha, de cabeçalho e hifenização foram eliminadas do texto. Foi preciso indicar na configuração do *WS* que os caracteres que servem de acento podiam estar no meio de palavras, para que as palavras acentuadas não fossem divididas e fossem contadas como unidades. Os caracteres de aspiração mais os acentos que antecedem uma vogal inicial, por exemplo, “, fizeram contar, como formas lexicais diferentes, as mesmas palavras, que ora aparecem no texto com iniciais maiúsculas, ora sobre minúsculas, no meio da sentença. Um exemplo claro é a separação da freqüência de *Edoxe*, “Εδοξε (início de sentença, “**Edoce**, em código beta) e *edoxe*, ἔδοξε (no meio da sentença, “**e doce**, em código beta). Essa diferenciação foi mantida aqui, sem se utilizar o recurso da lematização¹³. Esse procedimento foi adotado em um único arquivo contendo toda a obra ou o corpus de referência, do qual foi extraído o corpus do quinto livro, isto é, o corpus de estudo. Portanto, eventuais erros de

¹² *unusually frequent*, no original, M. Scott, 1997, “How key words are calculated”, WordSmith Tools Help.

¹³ Lematizar significa juntar, na lista de palavras, as entradas com o mesmo lema, desprezando as variações de número, pessoa, caso etc. Do grego *lêmma*, *lêmmatos* (*to*), o que é possuído.

contagem, em virtude do tratamento dos caracteres do código beta no WST, mantiveram-se constantes nos dois corpora.

Definição dos corpora. O quinto livro foi incluído no corpus de referência, de modo a garantir que palavras-chave, em pequena quantidade, não aparecessem nos resultados simplesmente por não existirem no corpus maior. Sabe-se, de um lado, que a diferença de tamanho entre o corpus de estudo e o de referência pode afetar a quantidade de palavras-chave resultantes nos corpora de língua inglesa; de outro, que os estudos apontam uma estabilização da quantidade de palavras-chave a partir de uma diferença de cinco vezes do tamanho do corpus de estudo (Berber Sardinha, 1999a). No entanto, como não há estudo similar para a língua grega antiga, não foi estabelecido nenhum critério nesse sentido. No corpus de referência, a ferramenta *Wordlist* contou 64.559 itens (*tokens*) e 12.894 formas (*types*). No corpus de estudo, contou 4.752 itens e 1.850 formas.

4. Resultados

Tabela 1. Levantamento automático de palavras-chave do Livro 5 em relação a toda obra da *Oneirocritica*, obtidas pela ferramenta Keywords.

Keyword	n-Livro 5	% Livro 5	N-Obra	% Obra	Keyness p≤.000001	P
Ἔδοξε	87	1.83	87	0.13	238.8	0.000000
τις	89	1.87	393	0.61	72.4	0.000000
ἦν	40	0.84	115	0.18	53.9	0.000000
αὐτῶ	54	1.14	209	0.32	52.4	0.000000
αὐτοῦ	37	0.78	110	0.17	48.2	0.000000
ἐγένετο	27	0.57	56	0.09	48.1	0.000000
τοῦ	92	1.94	533	0.83	47.1	0.000000
ἀπέθανεν	19	0.40	35	0.05	36.8	0.000000
τὸν	83	1.75	593	0.92	25.9	0.000000
εἰκότως	15	0.32	35	0.05	24.3	0.000001
ἦ	41	0.86	1.142	1.77	26.1	0.000000
καὶ	245	5.16	4.751	7.36	35.2	0.000000
τοῖς	9	0.19	693	1.07	50.7	0.000000
δε	87	1.83	2.591	4.01	69.1	0.000000

A tabela 1 mostra os resultados do levantamento de todas as palavras-chave da ferramenta *Keywords*: a lista de palavras, sua frequência absoluta e relativa no corpus de estudo (livro 5 da *Oneirocritica*), sua frequência absoluta e relativa no corpus de referência (todos os livros da *Oneirocritica*), o valor crítico de “chavicidade” (*keyness*) e a probabilidade associada a esse valor.

Tabela 2. Porcentagem, análise morfológica e sentido possível no texto das palavras-chave do livro 5 da *Oneirocritica*.

Palavras Livro 5	% na obra	Análise Morfológica	Sentidos possíveis no livro 5
Ἔδοξε	100	Verbo <i>dokeō</i> , Indicativo Aoristo, 3ª p. sing.	sonhou
τις	22	Pronome indefinido, fem/masc.	Um(a), algum, alguém
ἦν	34	Verbo <i>einai</i> , Indicativo, Imperfeito, 1ª/3ª p. sing.	era, estava
αὐτῷ	25	Pronome pessoal/dem. 3ª p., dativo, masc., sing.	para ele/ a esse, lhe
αὐτοῦ	33	Pronome pessoal 3ª pessoa, genitivo, masc., sing.	dele/seu; desse/ esse
ἐγένετο	48	Verbo <i>gignomai</i> , Indicativo, Aoristo, 3ª p. sing.	tomou-se; aconteceu para/ teve (+dat); ficou
τοῦ	17	Artigo/pron. genitivo, masc. sing.	do/ o
ἀπέθανεν	54	Verbo <i>apothnēskō</i> , 3ª p. sing. Indicativo, Aoristo	morreu
τὸν	14	Artigo/pron. masc. sing. acusativo	o
εἰκότως	42	Advérbio	justamente, naturalmente

Tabela 3. Frequência dos termos selecionados no Livro 5 e suas funções básicas no texto do Livro 5.

Palavras	Livro 5 n	Total N	Função no Livro 5
Ἔδοξε τις	84	84	introduzir o relato do processo mental do sonho (intercalado ou não com identificador/atributo do sujeito)
ἔδοξε (v)	10	145	relatar o processo do sonho
ἔδοξε τις	0	55	-
ἐγένετο	27	56	indicar evento(s) pós-sonho
ἀπέθανεν	19	35	indicar a ocorrência da morte
ἀπέθανε	8	16	"
ἀπέθανεν +			-
ἀπέθανε	27	51	
εἰκότως	15	37	modalizar o desfecho, introduzir explicação

As palavras-chave em **negrito** correspondem às que têm chavicidade positiva, e as em *itálico*, à chavicidade negativa. São palavras com freqüência alta, acima do esperado em comparação com o corpus de referência: **τις**, ἦν, αὐτῷ, αὐτοῦ, ἐγένετο, τοῦ, ἀπέθανεν, τὸν, εἰκότως, (*édoxe, tis, ên, autô, autoû, egéneto, tou, apéthanen, ton, eikótôs*), com a probabilidade associada ao acaso ser menor ou igual a 0,000001. A lista de palavras-chave positivas contém verbos, pronomes, artigos e advérbio. As palavras-chave de conteúdo são principalmente os verbos: Ἔδοξε (*édoxe*), sonhou, ἐγένετο (*egéneto*), tornou-se, e ἀπέθανεν (*apéthanen*), morreu (tabela 2). O verbo *ser*, artigos e pronomes, bem como as palavras-chave negativas, foram desprezados desta análise com exceção do pronome indefinido **τις**, por motivos discutidos adiante. Um comentário é feito sobre o advérbio *eikótôs* (εἰκότως).

Como não houve lematização, as tabelas 1 e 2 não incluem as variações de pessoa ou número. Por questões da localização da acentuação no código beta, a freqüência de *édoxe* no início da sentença foi computada separadamente desta quando no meio da sentença. Note-se que a forma inicial desse verbo no quinto livro representa 100% de ocorrência de toda a *Oneirocritica*. A mesma distinção na contagem ocorre com as variações do verbo: *apéthanen*, com *-n* (nu eufônico ou móvel), e sem o *-n*: *apéthane*. As duas formas somadas representam 52,9% das ocorrências dessa forma verbal no livro 5, em relação a toda a obra.

5. Análise dos resultados

Apéthane(n). O verbo morrer aparece no quinto livro como uma das possibilidades de realização de sonho, tanto para o sonhador quanto para pessoas relacionadas a ele no sonho. A figura 1 mostra os sonhadores e pessoas afetadas pela morte e sua relação com o sonhador. A morte é apresentada como um fato ocorrido no passado, após o sonho, cujo conteúdo, pela explicação do oniró-

crita, prenunciava esse resultado. A forma do Indicativo Aoristo fornece esse aspecto e modo verbais. A forma do verbo na 3ª do singular revela predominância de relato de prognóstico para uma única pessoa, com exceção de um caso de concordância com o sujeito no neutro plural: *trígêmeas (tríduma thêluká, ta tría. Onir. 5.12)*. Neste caso, a morte ocorreu para as filhas da sonhadora.

Eis alguns exemplos do desfecho *apêthane(n)*. Em um deles, a morte afetou o próprio sonhador:

Um indivíduo sonhou que se desprendia da carne como uma serpente da sua velha pele. No dia seguinte, morreu. Com efeito, a psique, estandø para abandonar o corpo lhe mostrava tais fantasias (Onir. 5.40).

Em outro, o sonhou previu a morte da irmã e o recebimento de uma herança:

Um indivíduo, que tinha uma irmã rica e ao mesmo tempo doente, sonhou que havia uma figueira diante da casa dela e que ele havia cozido e comido sete figos negros. A irmã morreu, tendo sobrevivido por sete dias, deixando o sonhador como herdeiro (Onir. 5.35).

Nesse caso, a morte atinge uma das filhas de quem sonhou:

Um pai, que tinha duas filhas solteiras, sonhou que uma tinha, sobre a cabeça, presa uma Afrodite de ouro e que, da outra, havia nascido um ramo de videira. A primeira se casou e a outra morreu (Onir. 5.39).

Artemidoro explica que a deusa simboliza as bodas conforme verso de Homero e a videira, a morte, pois nasce da terra onde os cadáveres se decompõem e é desprovida de fruto na ocasião de maior esplendor.

Nº REF. SONHADOR	QUEM MORREU	
5.13 Um garoto lutador	O sonhador	}
5.26 Um homem	O sonhador	
5.27 Um homem morando no exterior	O sonhador	
5.40 Um homem	O sonhador	
5.33 Um homem	O sonhador	
5.52 Um homem que tinha irmão	O sonhador	
5.59 Um homem	O sonhador	
5.60 Um homem cujo irmão era falecido	O sonhador	
5.62 Um homem	O sonhador	
5.93 Uma pessoa	A sonhadora	
5.63 Uma mulher	A sonhadora	
5.30 Uma mulher doente	O sonhador	
5.71 Um homem doente	O sonhador	
5.92 Um homem doente	O sonhador	
5.94 Um homem que estava para ser operado		
<hr/>		
5.25 Um homem (amante)	A amante do sonhador	}
3.35 Um homem (irmão)	A irmã do sonhador	
5.23 Um escravo	O mestre do sonhador	
5.85 Um escravo	A senhora do sonhador	
<hr/>		
5.12 Uma mulher que teve trigêmeas	As filhas trigêmeas da sonhadora	}
5.39 Um homem com 2 filhas solteiras	Uma das filhas do sonhador	
5.50 Um homem servindo como embaixador	O filho mais novo do sonhador	

Figura 1. Sonhadores e pessoas afetadas por *apêthanen* no livro 5.

Nº.REF	SONHADOR	ATRIBUTO ADQUIRIDO (tornou-se...)	PORTADOR DO ATRIBUTO
5.11	um homem,	Cego	τυφλὸς ἐγένετο
5.20	um senhor de escravo,		
5.44	um atleta,		
5.77	pai de atleta,		
5.90	um homem		
5.1	um homem	sacerdote	ἱερεὺς ἐγένετο
5.3	um homem	manco	χαλὸς ἐγένετο
5.54	um homem	caolho	μονόφταλμος ἐγένετο
5.56	um navegante	em grande perigo	ἐν κινδύνῳ ἐγένετο μεγάλῳ
5.65	um homem	notório depravado	ἀποπεφασμένος κιναιδὸς ἐγένετο
5.95	um atleta	vencedor no jogos	ἐγένετο ἱερνίκης
5.30	uma mulher doente	gravemente doente	ἐγένετο ἐν νόσῳ πάνυ κινδυνάδει
5.73	uma mulher doente	(nã) mãe	μητὴρ δὲ οὐκ ἐγένετο
5.61	um homem	teve um abscesso na barriga	τοῦτω ἀπόστημα ἐγένετο κατὰ τῆς γαστρὸς
5.67	um homem	teve um filho	ἐγένετο παῖς αὐτῷ
5.86	uma mulher	teve um filho	ἐγένετο αὐτῇ υἱός
5.74	um camponês	dono de navio	γεωργοῦ παῖς ναύκληρος ἐγένετο
5.76	um homem (pai de atleta)	vencedor nos jogos olímpicos	ὁ παῖς Ὀλυμπιονίκης
5.67	um homem que teve sexo com prostituta	defeituoso	παῖς ἐγένετο αὐτοῖς ἐπίμομος
5.57	um homem sem filhos	ilustre e famoso	ἐγένετο παῖς διάσημος καὶ λαμπρὸς
5.47	um homem	ilustre e famoso	ὁ υἱὸς ἐγένετο διάσημος καὶ λαμπρὸς
5.53	uma mulher	causa de injúrias e abusos	ζημιῶν καὶ πομπειῶν ἐγένετο αὐτῇ αἰτία.

Figura 2. Sonhadores e atributos marcados por *egéneto*

No livro 5, encontram-se apenas dois casos do verbo na forma do plural, *apéthanon* (morreram), referindo-se aos filhos do sonhador. Pode-se inferir que a presença da morte como prognóstico de boa parte dos exemplos de sonhos relatados no quinto livro manifesta a grande preocupação ou medo por parte do autor e, possivelmente, de clientes, homens, mulheres e escravos, com relação à própria morte, à de seus filhos, amantes, irmãos e senhores. Além da evidência lexical do verbo morrer, há, no mesmo livro, presença da morte como modos de morrer: ser condenado e decapitado (*Onir.* 5.6), afogar-se (*Onir.* 5.22) ou ser morto por alguém (*Onir.* 5.15). A lista de palavras-chave detectou uma forma específica de expressão do morrer na obra. Concordâncias do mesmo verbo em outras formas, como suas formas nominais e o substantivo correspondente, poderão ser testadas futuramente nos outros livros. Deve-se ressaltar que Artemidoro dedicou uma seção de seu segundo livro a sonhos com conteúdos de morte.

Egéneto. Apesar de o verbo *gígnomai* possuir sentidos cuja tradução nem sempre pode ser feita pela mesma palavra em português ou em outras línguas, a função de *egéneto* no quinto livro é apresentar algum evento ou situação que mudou ou passou a acontecer após o sonho. Como *apéthanen*, *egéneto* está na 3ª pessoa do singular do Indicativo Aoristo, referindo-se a um evento passado, relatado como fato. A figura 2 mostra os resultados dos sonhos associados a *egéneto*. Os atributos adquiridos após o sonho são, em geral, do próprio sonhador, mas há relatos em que a mudança existencial ou relacional refere-se ao filho. A integridade física também aparece como uma preocupação (*Onir.* 5.1; 5.3; 5.54). Um evento particularmente destacado é 'tornar-se cego' em cinco casos apontados pelo onirócrita (*Onir.* 5.11; 5.20; 5.44; 5.77 e 5.90). Nesses casos, os atributos são adquiridos pelo sonhador. A relação parental é notória, de um lado, em virtude de um dos atributos adquiridos ser a própria qualidade parental, ou seja, ter ou não um filho (*Onir.* 5.67; 5.86; 5.73) e, de outro, pelos atributos prognosticados no sonho referirem-se aos filhos dos sonhadores (*Onir.* 5.74; 5.57;



5.67; 5.74; 5.76). A preocupação com a situação profissional também se manifesta nos atributos adquiridos, tanto pelo sonhador quanto por seu filho: tornar-se dono de navio, ilustre e famoso na profissão, sacerdote e atleta vencedor. Duas preocupações singulares aparecem, uma relacionada aos costumes, à depravação e falta de pudor (*Onir.* 5.65) e outra em relação à preservação da família: a criada ser o pivô da separação de seus senhores (*Onir.* 5.53). Pomeroy (1991) observou que a *Oneirocritica* parece sustentar a idéia de família nuclear antiga, seguindo a tradição apontada por Paul Veyne (1987), dada a freqüência com que os familiares são mencionados. Ainda sugere que a preocupação com doenças e a mortalidade infantil revelam as condições de insalubridade, derivadas possivelmente da fome e falta de trabalho na época, e ambientes vividos pelo autor.

Com poucas variações de desfechos e sendo a morte um desfecho marcante, nota-se um contraste desses com a plasticidade criativa do autor nas variações de conteúdos oníricos. O sonho de um escravo com uma estrela que caía do céu e com outra que subia da terra para o céu (*Onir.* 5.23) pressagiou a morte do senhor e a posição assumida pelo filho na casa. O sonho de outro escravo (*Onir.* 5.85) em que recebera um ovo cozido de sua senhora, jogara a casca e comera o ovo, pressagiou a morte da senhora, que estava grávida, e os cuidados com a futura criança, que ficou a cargo do escravo. O sonho em que alguém acendeu uma lâmpada proveniente da lua pressagiou a cegueira (*Onir.* 5.11), porque a lua não tem luz própria e não poderia fornecer luz para acender qualquer coisa. O sonho de alguém em que carregava uma grande quantidade de ouro nos ombros também pressagiou a cegueira, porque a luz do ouro ofusca quem o carrega (*Onir.* 5.90).

Édoxe tis. Os relatos de sonho do livro 5 são iniciados, em sua maioria, por essa expressão. O pronome indefinido *tis* é o sujeito da ação de sonhar. Nesse sentido, esse pronome é uma marca da presença do sonhador. Ao empregá-lo, Artemidoro não queria dar

destaque aos sonhadores, mas ao sonho e seu resultado. Preferiu dar um caráter impessoal e indefinido ao sonhador: “um homem”, “uma mulher”, “um lutador”, “uma pessoa doente”, e assim por diante. Sendo *tis* uma marca padronizada de um tipo de presença do sonhador, as concordâncias automáticas dessa expressão auxiliaram a localização dos sonhadores. Em alguns casos, o atributo do sonhador foi encontrado explicitamente após o termo *tis*. Estão explícitos: um homem processado pela justiça (*Onir.* 5.10), uma mulher (*Onir.* 5.12; 5.37; 5.53; 5.86), um dono de navio (*Onir.* 5.16), uma mulher doente (*Onir.* 5.30), um atleta (*Onir.* 5.44), um lutador (*Onir.* 5.45), um homem doente (*Onir.* 5.71), um flautista (*Onir.* 5.81), um escravo (*Onir.* 5.85), um doente (*Onir.* 5.89), um atleta (*Onir.* 5.95). A expressão *édoxe tis* introduz o sonho em outros casos, nos quais fica implícito o atributo do sonhador pela declaração de uma relação, seja de parentesco ou social, de quem sonha com a pessoa que é mencionada no sonho em seguida. São implícitos: o ser marido (o que sonha com a própria mulher, *Onir.* 5.2), senhor (o que sonha com seu escravo; *Onir.* 5.20); pai (que sonha com o filho, *Onir.* 5.22; 5.41; o que tem duas filhas solteiras, *Onir.* 5.39; o que tem três filhos, *Onir.* 5.42; o que sonha levar o filha para lutar, *Onir.* 5.76); irmão (o alguém que tem uma irmã, *Onir.* 5.35; 5.52; o que sonha com o irmão preso a si, *Onir.* 5.60) e assim por diante. O sonhador é apresentado por Artemidoro no livro 5 por alguns atributos, relativos principalmente à posição familiar, gênero (homem ou mulher), ofício e estado de saúde.

O advérbio *eikótôs* pode significar ‘naturalmente’, ‘justamente’ ou ‘apropriadamente’, dependendo do contexto. Pela ferramenta *Concord* e a visualização da posição dos colocados (*collocates*) observou-se sua proximidade com a partícula explicativa *gar* (γάρ), que acompanha a proposição de explicação onirocrítica em 12 das 15 colocações de *eikótôs*. Trata-se de uma marca que introduz a explicação do desfecho do sonho. Não seria um advérbio que descreve o modo como esse desfecho ocorreu, mas indica a posição ou

opinião do autor acerca dos resultados; ele os considera esperados, evidentes e prováveis. Esse advérbio, portanto, poderia ser considerado uma marca de modalização ou da função interpessoal no texto, para usar a terminologia da teoria sistêmico-funcional (Halliday, 1984).

Conclusão

O emprego de uma ferramenta de análise lexical automática mostrou que é possível identificar algumas palavras-chave para sugerir aspectos relevantes acerca de preocupações do autor e temas no quinto livro da *Oneirocritica* (*apéthanen*, *egéneto*). Essas palavras-chave do quinto livro não se referem ao conteúdo onírico, mas a possíveis prognósticos, aqueles que foram relatados como acontecimentos reais, ocorridos após o sonho, reunidos a partir da experiência do onírico. Dentro desses acontecimentos, o termo *apéthanen* revelou a preocupação concreta e factual com a morte. As ocorrências de 'morreu' em relação ao sonhador são numerosas comparadas aos demais eventos. Pode-se dizer que as palavras-chave indicam duas classes principais de desfechos do sonho: a da morte (*apéthanen*) e da sobrevivência, ou transformação do indivíduo (*egéneto*) em seus atributos, que variam. Ambas as classes de desfechos atingem principalmente os sonhadores que são identificados pela posição familiar, gênero, ofício e estado de saúde, e sugerem as preocupações e anseios do autor e de sua possível clientela. A expressão inicial *edoxe tis* (Alguém/um(a)... sonhou) marca essa identificação do sonhador e o início do conteúdo onírico. O advérbio *eikótôs* parece ressaltar a posição do autor relativa ao relato do desfecho e a sua explicação onírica.

Não há evidências suficientes no quinto livro para se afirmar que Artemíodoro visava a uma clientela de elite, como sugere Bowersock. Os sonhos de escravos e mulheres poderiam ser um indício de uma clientela comum, não fosse pelo fato de que desfe-

chos do sonho, nos casos examinados, afetaram, de um lado, os padrões, com a morte desses, e, de outro, os maridos, com a dissolução familiar. É possível que os valores manifestados pela interpretação onirocrítica de Artemidoro, mesmo de sonhos ocorridos a pessoas comuns, sejam os de uma elite. Por outro lado, os problemas e preocupações relatados no quinto livro não parecem ser de uma elite, mas parecem estar associados a expectativas amplas, não exclusivas a classes: ter filhos, ter saúde, manter a integridade da família, dos bens, evitar perigos, evitar processos, não se tornar escravo, ser um profissional famoso, vencer nos jogos e, principalmente, manter-se vivo.

Para o levantamento inicial da morte como uma fonte de medo e preocupação na *Oneirocritica*, curiosamente não foi necessária uma análise dos conteúdos oníricos. A ênfase nos indícios de morte e nas poucas variações de desfechos, indicativos de mudanças do sonhador, contrastam com a criatividade e a plasticidade dos sonhos. A forte presença do verbo *sonhou* seguido do conteúdo do sonho ressalta essa variação. Em termos do *logos* ou discurso onirocrítico, o autor parece ter sido muito mais criativo nos sonhos e na explicação do que nos desfechos. O sonho representa a ficção, enquanto o presságio, a realidade, em que a expectativa de morte parece ocupar um grande espaço. Pode-se dizer que o foco do quinto livro não reside na elaboração onírica, mas em uma “ficção da realidade”, que acompanha o sonho, revelando, justamente, as preocupações e ansiedades do autor.

Este estudo, embora inicial, cujo objetivo foi exemplificar a utilização de um software que realiza análises lexicais baseadas em cálculos de frequência, mostra um resultado promissor do trabalho com tecnologia de informação aplicada a textos em língua grega antiga. Há muito por fazer. Alguns caminhos são sugeridos para futuros estudos da própria *Oneirocritica*. Do ponto de vista metodológico, poder-se-ia diminuir a restrição das palavras-chave do quinto livro, isolando o corpus de estudo do corpus de referência. Do pon-

to de vista temático, sugere-se, também, uma análise dos conteúdos oníricos, associados a outros medos e expectativas nos demais livros, por meio, por exemplo, da concordância de termos relacionados a essas afecções, ou da interpretação de termos de chavidade negativa no quinto livro ou adotando toda a *Oneirocritica* como o corpus de estudo, comparando-a a um corpus de referência composto de obras diversas. Para aprofundar a questão apresentada por Bowersock, sobre a diferença das preocupações manifestadas nos sonhos em romances gregos, seria possível realizar futuramente um estudo específico de gêneros, comparando-se a *Oneirocritica* aos romances. Para tal, o desenvolvimento de um sistema automatizado de recuperação de obras dentro dos arquivos de autores seria bem-vindo, permitindo formar, com facilidade, a partir do TLG, sub-corpora compostos de obras de diferentes autores, mas classificados dentro do mesmo gênero.

ABSTRACT: *This paper describes the use of WordSmith Tools, a software application for automated lexical analysis, in a study of the meaning of the pressages in onirocritical logos found in Artemidorus' Oneirocritica. The corpus studied is the fifth and last book of Oneirocritica and the reference corpus comprises all five books. The keywords obtained from the lexical analysis over the two corpora – apéthanen [died] and egêneto [became] – raised the hypothesis that death might be one of the main onirocritical predictions and concerns. A thematic search, performed with the tool, is demonstrated and its possibilities discussed.*

KEYWORDS: *Oneirocritica, Artemidorus, onirocriticism, lexical analysis, WordSmith Tools, Thesaurus Linguae Graecae, TLG Workplace, Greek literature.*

Bibliografia

- ARTEMIDORI DALDIANI (1963) *Onirocriticon*. Roger Pack (Ed.). Leipzig: Teubner.
- ARTEMIDORO (1975) *Il libro dei sogni*. Trad., introd. e note di Dario Del Corno. Milano: Tascabili Bompiani.
- ARTEMIDORO (1989) *La interpretación de los sueños*. Introducción, traducción y notas por Elisa Ruiz Garcia. Madrid: Gredos.
- ARTEMIDORUS (1990) *The interpretation of dreams*. Oneirocritica by Artemidorus. Translation and commentary by R. J. WHITE. Torrance, CA: Original Books.
- ARTÉMIDORE (1975) *La clef des songes*. Onirocriticon. Traduit et annoté par A. J. Festugière. Paris: Vrin.
- ARTÉMIDORE D'ÉPHÈSE (1964) *Le livre des songes*. Traduit du grec en arabe par Hunayn B. Ishaq. Fahd (Ed.). Paris: Damas.
- BERBER SARDINHA, T. (1999a) "A influência do corpus de referência na obtenção de palavras-chave". In: *Direct papers*, 38, LAEL/PUC-SP e AELSU, U. of Liverpool. <http://www.direct.f2s.com/DirectPapers38.pdf>. (19.set.00).
- _____. (1999b) Usando WordSmith Tools na investigação da linguagem. *Direct papers*, 40, LAEL/PUC-SP. <http://www.direct.f2s.com/Direct-Papers40.pdf>. (19.set.00).
- BERKOWITZ, L.; SQUITIER, K. A. (1990) *Canon of Greek author and works*. Thesaurus Linguae Graecae, 3rd edition. New York/Oxford: Oxford University Press.
- BIBER, D.; CONRAD, S.; REPPEN, R. (1998) *Corpus linguistics*. Investigating language structure and use. Cambridge: Cambridge University Press.
- BOWERSOCK, G. W. (1994) *Fiction as history*. Berkeley: University of California Press.
- CLAYMAN, D.; CRANE, G.; GUTHRIE, D. (1992) Response: Clayman, Crane, Guthrie on Keyser on Frischer. *Bryn Mawr classical review*, 3. <http://ccat.sas.upenn.edu/bmcr/1992/03.06.20.html> (19.set.00).
- COLLINS, H.; SCOTT, M. (1997) "Lexical landscaping in business meetings". In: BARGIELA-CHIAPPINI, F.; HARRIS, S. (Org.). *The languages of business - An international perspective*. Edinburgh: Edinburgh University Press.

- CRANE, G. R. (Ed.) (1997) *The perseus project*. <http://www.perseus.tufts.edu> (19.set.00).
- DUNNING, T. (1994) "Accurate methods for the statistics of surprise and coincidence". In: ARMSTRONG, S. (Ed.). *Using large corpora*. Cambridge, MA: MIT Press.
- FOUCAULT, M. (1984) *História da sexualidade, 3: o cuidado de si*. Trad. de Maria T. C. Albuquerque. Rio de Janeiro: Graal.
- FREITAS, A. C. de. (1997) *América mágica, Grã-Bretanha real e Brasil tropical: um estudo lexical de panfletos de hotéis*. Tese de Doutorado, LAEL/ PUC-SP. São Paulo. http://www.lael.f2s.com/tese_alice.zip (19.set.00).
- HALLIDAY, M. (1994) *An introduction to functional grammar*. London: Edward Arnold.
- HOULIHAN, J. (1997) "Observations on the text of Artemidorus". In: *Illinois classical studies*, Onirocritica Book 1. 22: 99-119.
- KEYSER, P. (1991) "Rev. of *Re-counting Plato: a computer analysis of Plato's style*, por G. R. Ledger". In: *Bryn Mawr classical review*, 2. <http://ccat.sas.upenn.edu/bmcr/1991/02.07.03.html> (19.set.00).
- LEDGER, G. R. (1989) *Re-counting Plato: a computer analysis of Plato's style*. Oxford: Clarendon Press.
- LEDGER, G. R.; KEYSER, P. (1992) "Responses (two): Ledger on Keyser; Keyser on Nails on Keyser on Ledger". In: *Bryn Mawr classical review*, 3. <http://ccat.sas.upenn.edu/bmcr/1992/03.06.19.html> (19.set.00).
- McENERY, T.; WILSON, A. (1996) *Corpus linguistics*. Edinburgh: Edinburg University Press.
- OAKES, M. P. (1998) *Statistics for corpus linguistics*. Edinburgh: Edinburgh University Press.
- POMEROY, A. J. (1991) "Status and status-concern in the greco-roman dream-books". In: *Ancient society*, 22: 51-74.
- RAMOS, R. G. (1997) *Projeção de imagem através de escolhas lingüísticas: um estudo no contexto empresarial*. Tese de Doutorado, LAEL/PUC-SP.
- RYDBERG-COX (2000) "Co-occurrence patterns and lexical acquisition in ancient Greek texts". In: *Literary and linguistic computing*, 15 (2): 121-30.
- SCOTT, M. (1997) *WordSmith Tools, v.2.0*. Software. Oxford: Oxford University Press.

TLG Thesaurus Language Graecae v. E. CD ROM. (2000). University of California: Irvine.

TLG Workplace v. 6.01. Software para leitura e recuperação de textos do *TLG*, para PC/Win95. (1997) Silver Mountain Software. <http://www.silvermnt.com> (17.jun.97).

_____. *v.7.01.* Software para leitura e recuperação de textos do *TLG*, para PC/Win95. (1999) Silver Mountain Software. <http://www.silvermnt.com> (11.nov.99).

_____. *v.8.0.* Software para leitura e recuperação de textos do *TLG*, para PC/Win95. (2000). Silver Mountain Software, <http://www.silvermnt.com> (22.fev.00).

VEYNE, P. (1987) *The roman empire*. Cambridge, MA: Harvard University Press.

WINKLER, J. J. (1990) *The constraints of desire*. The anthropology of sex and gender in ancient Greece. New York: Routledge.