

Contos da tradição oral: estudo informático-lexical e simbólico¹

Isabel Moreira

Pré-escolar de Picão – Castro Daire

Palavras-chave: conto, imaginário, linguagem, programas de tratamento automático de textos.

Keywords: tale, imaginary, language, automatic text processing.

1. A importância do léxico no desenvolvimento do imaginário

A viagem pelo imaginário começa a ser percorrida e desenvolvida desde tenra idade. Esse percurso, que está indubitavelmente relacionado com os contos da tradição oral, necessita de estímulos para que os efeitos desejados se façam sentir e o seu poder seja consolidado sob a forma de mentes mais abertas a uma efectiva participação activa enquanto seres humanos/sociais. Mentes que, por não terem sido privadas de estímulos essenciais, não perdem nunca a capacidade de «ficar perplexas». A narração de um conto é um manancial de puro prazer e exercício imaginativo, cujos benefícios serão sentidos sob a forma de uma sociedade mais justa e efectivamente humana, onde os valores essenciais à sua estruturação e desenvolvimento se mantêm fortes e inabaláveis.

O pedido de contos por parte da criança não é um pedido incipiente, bem pelo contrário: representa, mais do que uma demanda de conforto, uma demanda de evasão. É uma solicitação plena de anseios por uma realização imaginativa que lhe permitirá sair do seu eu real para penetrar num mundo onde tudo é possível, num mundo de que o

¹ Neste artigo, é apresentada uma síntese da Dissertação de Mestrado intitulada *Contos da Tradição Oral: Estudo Informático-lexical e Simbólico*, apresentada à Universidade de Trás-os-Montes e Alto Douro, em Setembro de 2006.

seu eu imaginativo necessita e onde a liberdade criativa, de que se revela capaz, a conduz a realidades inacessíveis no seu mundo real. Desta confrontação entre o eu real e o eu imaginativo surge a capacidade de lidar e ultrapassar os problemas com que a criança se vai deparando no seu processo evolutivo enquanto ser humano. Os perigos da negação do uso desta liberdade imaginativa resultam em frustrações dificilmente ultrapassáveis. Uma criança a quem é vedado o direito de exteriorizar as suas manifestações imaginativas, a sua liberdade criativa, e a quem é vedado o acesso a imagens propiciadoras de tais situações acabará por desenvolver receios e defesas contra este mundo exterior a ela e, inclusive, consciente ou inconscientemente, recusar-se a fazê-lo de novo. As consequências serão tremendas.

Para podermos compreender melhor o alcance que os contos têm junto das crianças e, como consequência, o quanto elas são negativamente afectadas pela falta deles, temos de em primeiro lugar saber que a criança tem um poder imaginativo, inato, sem limites. Ela cria e recria as realidades, vividas ou imaginadas, fazendo apelo a uma enorme capacidade de ultrapassar limites e barreiras, intransponíveis para quem, nas palavras de Erich Fromm, perdeu a «capacidade de ficar perplexo». A criança, ao contrário dos adultos «desprovidos de imaginação», tem a capacidade de «conceber mais do que meras duplicações dos acontecimentos e tramas da experiência concreta» (Fromm, 1980: 13). Ela consegue imaginar para lá do imaginável, mas esta capacidade inata de poder imaginativo não perdura no tempo se não for alimentada e estimulada. Ela precisa de ser confrontada com imagens, com sons, com símbolos, com contos...

Convictos da importância que o desenvolvimento da linguagem e a construção do imaginário na criança detêm no seu desenvolvimento e formação enquanto ser adulto, bem como da importância que esse mesmo imaginário tem na formação de valores essenciais à estruturação e desenvolvimento de uma sociedade, e certos de que a literatura tradicional, mais concretamente os contos, proporcionam estes pressupostos fundamentais, com *Contos da Tradição Oral: Estudo Informático-lexical e Simbólico*, pretendemos, fundamentalmente, aferir do léxico presente na literatura de tradição oral, mais precisamente nos contos, que se reveste de tal importância.

Com o propósito de descobrir as palavras que investem os contos de tal poder, trataram-se, em ambientes de desenvolvimento linguístico (com os programas de tratamento automático de textos – *NooJ* e *Lexicon*) quatro recolhas efectuadas no âmbito dos contos da tradição oral. Constituem o corpus deste estudo as recolhas efectuadas por Francisco Adolfo Coelho (1879), Teófilo Braga (1883-1914/15), Consiglieri Pedroso (1910) e ainda uma recolha efectuada em Trás-os-Montes, por Joaquim Alves Ferreira, e publicada já em finais do século XX, mais precisamente em 1999.

As conclusões retiradas desse tratamento, apresentadas sob a forma de resultados sistematizados e rigorosamente fiáveis, foram diversas e algumas surpreendentes.

Contudo, uma coisa fica clara, o poder dos contos da tradição oral é ilimitado: os contos representam o mais puro e verdadeiro imaginário da alma humana. São elementos identitários da mais fundamental importância e cujo alcance pedagógico não tem limites. E, se tempos houve em que o seu interesse se limitava à sua colecta, estudo das origens e do conteúdo, hoje em dia ninguém duvida do seu poder enquanto elemento essencial na formação e desenvolvimento da criança e, portanto, presença indispensável na sua vida.

2. O estudo dos contos da tradição oral

Desde que Platão, em *A República*, no capítulo dedicado à educação, atribuiu aos contos uma importância fundamental para educar os «homens em imaginação», referindo a necessidade de persuadir «as mães e as crianças a contá-las² às crianças, e a moldar as suas almas por meio das fábulas, com muito mais cuidado do que os corpos com as mãos» (Platão, 2001: 86 e 87), ficou bem marcada a importância desta forma breve do género narrativo, que viria a ser objecto de estudo de várias disciplinas científicas.

Com efeito, ao longo dos anos, os contos da tradição oral têm sido objecto de estudo de disciplinas como a História, a Etnologia, o Folclore, a Literatura, a Linguística, a Psicologia e a Psicoterapia.

Os contos da tradição oral, cujo garante de sobrevivência residiu na sua capacidade de serem memorizados, são hoje, depois de terem sido objecto de estudo sistemático, principalmente a partir do século XIX, juntamente com outras manifestações da chamada literatura tradicional oral, considerados, enquanto manifestações de um património cultural colectivo anónimo, elemento identitário que transporta consigo um tipo de imaginário individual e colectivo reflectido na realidade simbólica de que se reveste, «elemento importante para o conhecimento da nossa cultura encarada no contexto mais vasto de uma cultura ancestral» (Pedroso, 2000: 28).

A origem dos contos, sempre denominada de «remota», está ainda por determinar. Adolfo Coelho faz alusão a que «os contos em grande parte remontam, não só à Idade Média, mas mesmo à Antiguidade» (Coelho, 2005: 27). Atendendo a que Platão, como já vimos, lhes faz referência, no seu tratado *A República*, como sendo elemento fundamental na educação e atendendo, ainda, à referência a Homero «divinizado pela Grécia» por ser um «sublime contador dos contos da carochinha» (Queirós, 2000: 208)³,

² Fábulas é o termo que o autor utiliza, numa acepção genérica de narrativa curta.

³ A citação faz parte de uma carta escrita pelo autor aos Condes de Arno e de Sabugosa, incluída em *Correspondência* (páginas da vida íntima e literária), datada de 8 de Fevereiro de 1895.

não nos é difícil atribuir aos contos uma origem que, ainda por determinar, seja com efeito de uma época muito remota.

Georges Jean, no prefácio de *O Fio da Memória*, reforça esta ideia quando diz:

E acontece que desde a noite dos tempos, os homens contam. Sem recorrer à escrita, mas transmitindo de boca em boca como o aedo helénico, histórias cheias de aventuras, de personagens simples e de coisas maravilhosas. Estas histórias que constituem o «tesouro dos contos» e que se encontram em todas as culturas e sob formas infinitamente diversificadas tornaram-se para nós [...] a nossa memória colectiva em que se exprimem todos os matizes que fazem a originalidade, mas também a universalidade das nossas culturas nacionais (Traça, 1998: 7).

A título de curiosidade, refira-se que o conto mais antigo que se conhece é «Os dois irmãos», conservado num papiro egípcio do século XIII A.C. (história de dois irmãos, Anepu e Batau, um que fica em casa e outro que vai pelo mundo. O que fica em casa vai ser avisado por um objecto mágico de que o seu irmão corre perigo de morte, indo-o buscar e salvando-o).

Se a origem remota dos contos, bem como a sua formação e transmissão ao longo dos séculos sem sofrerem grandes transformações, continua a constituir um enigma, que dificilmente se decifrárá (segundo alguns antropólogos, a origem dos contos remonta a práticas religiosas e a rituais que também deram origem a mitos), não o é menos a existência de semelhanças flagrantes entre contos de todo o mundo. Têm sido estes, de resto, os factores aos quais se atribui o fascínio exercido sobre inúmeros estudiosos das mais diversas áreas, sistematicamente desde o século XIX, e que levou a que os contos se tenham constituído como seu principal objecto de estudo.

O conto, ainda que «tão antigo como o próprio homem» (Ferreira, 1999: 3), começou a despertar um interesse diferente durante o século XVII. Se podemos dizer que antes deste período os contos tinham já valor reconhecido (factor que decididamente o fez atravessar povos e gerações tendo como único suporte a memória), foi neste século XVII, onde os «Contos de Fadas» se constituíram como um género muito em voga, que eles começaram a ganhar um estatuto diferente do simples reconhecimento de valor ou do interesse pelo prazer que provocavam.

Durante este século, aparecem vários livros do género, de entre os quais destacamos o de Charles Perrault, de 1697, intitulado *Histoires ou Contes du Temps Passé*. Ele, o seu filho, ou ambos, adaptaram com grande mestria uma série de contos tradicionais franceses, destinados à leitura da infância, já que, e segundo o autor, citado por Adolfo Coelho, «eram histórias que todas as velhas criadas contavam às crianças, ao serão» (Coelho, 2005: 16). Ainda que o valor destes contos tenha passado despercebido nesta altura, foi a partir deles, e depois definitivamente com os Irmãos Grimm, que os contos da tradição oral passaram a ser considerados contos para a

infância, carácter que os próprios Irmãos Grimm lhes atribuíram ao intitular a sua recolha de *Kinder – und Hausmärchen* (contos para crianças e para o lar). É também nesta altura que o destino da literatura tradicional escrita, e adaptada, vai mudar dentro das classes letradas, ainda que não no meio do povo, já que continua a ser um género muito «popular» ao nível, se não da produção, da recepção.

Foi definitivamente no século XIX, mais concretamente em 1812 e 1814, quando os Irmãos Jacob e Wilhelm Grimm concentraram esforços no estudo da língua alemã, numa luta romântica em prol do povo e da nação, que os contos da tradição oral despertaram o interesse e curiosidade científicos, que fariam deles objecto de estudo sistemático nos mais variados ramos das Ciências Sociais. Fundadores da filologia alemã, os irmãos Grimm entregaram-se de corpo e alma aos estudos do alemão arcaico, à procura das raízes esquecidas. Estudos comparativos dos vários dialectos e uma recolha etnográfica dos contos viriam a revelar, para além de valores linguísticos (os estudos revelaram que os vários dialectos eram variantes de uma língua comum, que fazia dos povos alemães uma só raça), valores culturais, crenças e costumes que viriam, posteriormente, com o surgimento da mitologia comparada, a enraizar-se como mitos universais.

No que se refere a Portugal, a primeira colectânea a ser publicada, 65 anos passados sobre a dos irmãos Grimm, seria a de Adolfo Coelho, em 1879. A esta colectânea segue-se a de Teófilo Braga, em dois tomos, datada de 1883 e 1914/1915, e a de Consiglieri Pedroso, em 1910. Os trabalhos destes três autores, precisamente os que, juntamente com o trabalho mais recente de Joaquim Ferreira, constituem objecto do nosso estudo, já haviam sido, contudo, precedidos por outros estudos dos mesmos autores no âmbito dos contos, e inclusive, em alguns casos, publicados em revistas especializadas, como por exemplo a recolha efectuada e publicada por Consiglieri Pedroso que, embora datada de 1910, foi precedida de outra, parcial, em inglês, editada pela Folk-Lore Society, de Londres, em 1882.

3. Os contos e o desenvolvimento da criança

Bruno Bettelheim foi quem mais se destacou no estudo dos contos de fadas sob o ponto de vista da sua importância no desenvolvimento psicológico da criança. Para este autor, a ideia errada que, mais generalizadamente do que seria desejável, se tenta transmitir à criança de que o mal não é inerente ou não existe na natureza humana, de que todos os homens são bons, é um primeiro e imensíssimo erro cujas consequências se podem revelar da mais extrema gravidade no desenvolvimento da criança, colocando em risco todo o seu desenvolvimento, «penhorando» a sua oportunidade de realização e formação como ser adulto.

Segundo o autor, a não confrontação, desde tenra idade, da criança com o «lado

sombrio» que existe na natureza humana, havendo, pelo contrário, uma tendência a esconder da criança este facto, fazendo-a acreditar que todos os seres humanos são bons, vai provocar, inevitavelmente, uma confrontação com a realidade, a dela própria enquanto ser humano com todas as limitações e imperfeições a que está sujeito, que se revelará em conflitos e frustrações dificilmente ultrapassáveis. Os contos de fadas, acepção que o autor dá aos contos, não só evitam que isto aconteça, como ainda preparam a criança para ser capaz de dar respostas aos conflitos que ela mesma experimentará enquanto ser humano. Nas palavras do autor,

Há uma recusa generalizada em deixar as crianças saberem que a fonte de muito do que vai mal no mundo é devido às nossas próprias naturezas – a propensão que todo o homem tem para agir agressivamente, associalmente, egoistamente, por raiva ou angústia. Em vez disso, queremos que os nossos filhos acreditem que todos os homens são bons por natureza. Mas os miúdos sabem que *eles* não são sempre bons; e muitas vezes, mesmo quando o são, prefeririam não o ser. Isto vem contradizer o que os pais lhes dizem, o que faz com que a criança se veja a si própria como um monstro (Bettelheim, 2005: 15).

O papel dos contos vai, neste contexto, ser fundamental, preparando a criança para uma realidade que é, afinal, a dela, e mostrando-lhe que não há problemas sem solução, que todos os conflitos que ela inevitavelmente irá sentir têm uma forma de ser ultrapassados e vencidos. «É esta precisamente a mensagem que os contos de fadas trazem à criança, por múltiplas formas: que a luta contra graves dificuldades na vida é inevitável, faz parte intrínseca da existência humana – mas que, se o homem se não furtar a ela, e com coragem e determinação enfrentar dificuldades, muitas vezes inesperadas e injustas, acabará por dominar todos os obstáculos e sair vitorioso» (ibid.: 15).

De entre os autores que são objecto do nosso estudo, o que primeiro se referiu ao conto como portador de uma função pedagógica foi Adolfo Coelho. Ernesto Veiga da Silva, em relação ao papel que Adolfo Coelho desempenhou neste campo, refere que «a tónica fundamental deste sector da sua obra é a importância que atribui às relações entre a Psicologia ou Educação e a Etnografia, ou, mais concretamente, ao papel essencial que cabe aos elementos tradicionais na educação, nomeadamente o conto popular» (Coelho, 2005: 32). Teófilo Braga, para quem os contos desempenham também um papel fundamental na educação da criança, refere que «os contos tradicionais são imensamente simpáticos às crianças» e que, portanto, «a importância dos contos tradicionais é evidente para a educação das crianças» (Braga, 2002: 22 e 23).

Podemos dizer que, dos muitos aspectos que se relacionam com o estudo dos contos e sobre os quais se pode, em muitos casos, não chegar a um consenso, há um dado sobre o qual não existe qualquer dúvida e em que as opiniões são, hoje, unânimes:

o poder dos contos da tradição oral na educação é inegável.

Um dos primeiros pedidos (e depois o mais frequente), que qualquer criança em idade pré-escolar faz é «Conta-me uma história...». E o mágico «Era uma vez» é a resposta, ansiosamente esperada, a este pedido. Ao longo dos últimos 13 anos, em contacto directo com crianças em idade pré-escolar, assumindo também a função (privilegiada) de contadora de histórias, apercebi-me de que o poder dos contos não tem limites e que nada se compara ao conforto, segurança e poder imaginativo que é transmitido pelo mágico «Era uma vez...», pois, como defende Bettelheim, «Aplicando o modelo psicanalítico da personalidade humana, os contos de fadas são portadores de mensagens importantes para o psiquismo consciente, pré-consciente ou inconsciente, qualquer que seja o nível em que funcione. Lidando com problemas humanos universais, especialmente com os que preocupam o espírito da criança, as histórias falam ao seu *ego* nascente, encorajando o seu desenvolvimento, enquanto, ao mesmo tempo, aliviam tensões pré-conscientes ou inconscientes» (Bettelheim, 2005: 12 e 13).

Num grupo de crianças, espectadoras atentas e ansiosas pela expressão mágica do «Era uma vez», ansiosas pela história desejada, cada movimento, cada entoação, cada palavra pronunciada é fonte de genuíno contentamento e origem de experiências imaginativas, cujo alcance, incomparável, não se obtém se a ausência de histórias marcar presença. A narração de contos da tradição oral é, de facto, fundamental nesta etapa de desenvolvimento. A ausência desta forma de narrativa numa fase crucial do desenvolvimento da criança e de despertar social pode provocar lacunas ao nível afectivo, social e emocional, cuja falta se notará em todas as áreas de conteúdo e, logo, aos níveis educacional, social e familiar.

O conto da tradição oral estimula e favorece o desenvolvimento da criança ao nível das três grandes áreas de conteúdo: área de formação pessoal e social, área de expressão/comunicação⁴ e área de conhecimento do mundo.

Pela expressão oral, levantam-se questões sociais, lida-se com o comportamento humano e com etapas fundamentais da vida.

4. O carácter simbólico dos contos

O carácter simbólico dos contos, isto é, a comunicação simbólica que transmitem, confere-lhes o poder de se revestirem das mais variadas interpretações, de satisfazerem as mais variadas aspirações, os mais variados sonhos e desejos e, até, de darem resposta a medos e anseios. Pela identificação com as personagens, com as situações, com as

⁴ Que compreende três domínios: domínio das expressões, domínio da linguagem e domínio da matemática.

atitudes tomadas pelos «heróis» dos contos, as crianças vivem uma realidade simbólica que funciona como uma representação mental da sua estrutura interior, onde elas se projectam e se apoiam, dando forma e realizando o irrealizável. Os contos «falam-nos em linguagem simbólica», no conto «materializam-se simbolicamente os fenómenos psicológicos interiores» (Bettelheim, 2005: 49 e 50).

Erich Fromm, sobre esta linguagem simbólica, em *A Linguagem Esquecida*, refere que

A linguagem simbólica é uma língua em que as experiências íntimas, os sentimentos e pensamentos são expressos como se fossem experiências sensoriais, factos do mundo exterior. É uma linguagem cuja lógica difere da linguagem convencional que falamos de dia, uma lógica em que as categorias dominantes não são o espaço e o tempo, mas sim a intensidade e a associação. É o único idioma universal jamais criado pela raça humana, o mesmo para todas as criaturas e para todo o curso da história. É uma língua com gramática e sintaxe próprias [...] e cujo conhecimento é imprescindível para se poder entender o significado dos mitos, dos contos de fadas e dos sonhos (Fromm, 1980: 16).

O autor resume esta ideia ao dizer que «os símbolos [...] são imagens pictóricas ou palavras que substituem uma ideia, um sentimento ou um pensamento» (ibid.: 176). Estas imagens, por seu lado, satisfazem a necessidade vital de clarificar os processos inconscientes na criança. Bruno Bettelheim é claro ao afirmar que «Os processos inconscientes da criança só podem ser claros para ela através de imagens que falem directamente ao seu inconsciente. As imagens evocadas pelos contos de fadas fazem isso» (Bettelheim, 2005: 43). A eficácia dos contos de fadas reside, precisamente, no facto de que «nos contos de fadas materializam-se simbolicamente os fenómenos interiores» (ibid.: 50).

A questão relativa à simbologia está directamente relacionada com a do imaginário. Gilbert Durand, um dos mais conceituados investigadores no campo do simbolismo, diz-nos isso mesmo quando se refere à linguagem simbólica como pertencendo aos termos relacionados com o imaginário. A linguagem simbólica realiza-se, assim, no e pelo imaginário. Na tentativa de esclarecer aquilo que ele chamou de «extrema confusão na utilização dos termos relativos ao imaginário», estabelece a distinção entre dois tipos de signos a que chama de «*signos arbitrários*» e «*signos alegóricos*». Só depois de feita esta distinção poderemos chegar, nas palavras do autor, à «imaginação simbólica propriamente dita» e à sua abrangência. Segundo o autor

Podemos [...] distinguir entre dois tipos de signos: os *signos arbitrários* puramente indicativos, que remetem para uma realidade significada, se não presente pelo menos sempre apresentável, e os signos alegóricos, que remetem para uma realidade significada dificilmente apresentável. Estes últimos signos são obrigados a *figurar* concretamente uma parte da realidade que significam.

Finalmente, chegamos à imaginação simbólica propriamente dita quando o significado não é *de modo algum apresentável* e o signo só pode referir-se a um *sentido* e não a uma coisa sensível (Durand, 1993: 7-10).

O símbolo é, segundo o autor, que cita A. Lalande, «qualquer signo concreto que evoca, através de uma relação natural, algo de ausente ou impossível de perceber». O símbolo seria mesmo, citando P. Godet, o inverso da alegoria, já que esta «parte de uma ideia (abstracta) para chegar a uma figura, enquanto o símbolo é primeiro e em si figura, e, como tal, fonte, entre outras coisas, de ideias». O símbolo representa uma expressão mental espontânea e abstracta. Por outro lado, Gilbert Durand aproxima o símbolo da alegoria, mas alarga de forma concludente o seu domínio ao conferir-lhe a propriedade de «epifania», de «indizível».

o símbolo é, como a alegoria, recondução do sensível, do figurado ao significado, mas é também, pela própria natureza do significado inacessível, *epifania*, isto é, aparição, através do e no significante, do indizível [...]. O domínio de predilecção do simbolismo: o não sensível sob todas as suas formas: inconsciente, metafísico, sobrenatural e surreal. Estas «coisas ausentes ou impossíveis de perceber», por definição, vão ser, de maneira privilegiada, os próprios sujeitos da metafísica, da arte, da religião, da magia: causa primeira, fim último, «finalidade sem fim», alma, espíritos, deuses, etc. [...] Não podendo figurar a infigurável transcendência, a imagem simbólica é *transfiguração* de uma representação concreta através de um sentido para sempre abstracto. O símbolo é, pois, uma representação que faz *aparecer* um sentido secreto, é a *epifania* de um mistério (ibid.: 10-12).

Dando continuidade, e servindo-se das palavras de Paul Ricoer, refere o símbolo como possuidor de três «dimensões concretas». O símbolo, de uma abrangência maior que a sua metade «visível», o significante,

é simultaneamente «cósmico» (isto é, recolhe às mãos cheias a sua figuração no mundo bem visível que nos rodeia), «onírica» (isto é, enraíza-se nas recordações, nos gestos que emergem nos nossos sonhos e constituem, como bem demonstrou Freud, a massa muito concreta da nossa biografia mais íntima) e, finalmente, «poética», isto é, o símbolo apela igualmente à linguagem, e à linguagem que mais brota, logo, mais concreta. Mas também a outra metade do símbolo, a parte de invisível e de indizível que faz dela um mundo de representações indirectas, de signos alegóricos sempre inadequados, constitui uma espécie lógica bem à parte. Enquanto num simples signo o significado é limitado e o significante, ainda que arbitrário, é infinito: enquanto a simples alegoria traduz um significado finito por um significante e não menos delimitado, os dois termos de *Symbolon* são, por sua vez, infinitamente abertos. O termo significante, o único concretamente conhecido, remete em «extensão» [...] para

todas as espécies de «qualidades» não figuráveis, e isto até à antinomia. (ibid.: 12).

Os símbolos revestem-se, portanto, de tantas interpretações, e tão diferentes, que não se torna possível identificá-las e registá-las na sua totalidade e de forma concludente. O seu campo abrangente é tão vasto e diversificado quanto vasta e diversificada é a manifestação do espírito individual, pelo uso de uma liberdade de expressão criativa única em cada ser humano. Assim, por exemplo, uma *borboleta* para uma criança poderá significar a delicadeza, para outra a possibilidade de evasão ou a liberdade, para outra ainda poderá ser a imagem da leveza, para outra a da fragilidade, etc. Pode assumir, este como qualquer outro símbolo, uma significação tão vasta e diversificada que não é possível fazer mais do que tentar uma aproximação a algumas das suas possibilidades interpretativas.

As crianças simbolizam os seus desejos, representam simbolicamente o que de mais profundo existe no seu íntimo, remetendo as suas manifestações emocionais para uma realidade que, não sendo a delas, no sentido concreto e experiencial, o é, no contexto imaginativo e transcendente. Esta é, assim, uma «realidade» ao alcance delas e das suas necessidades inerentes e imediatas, manifestadas ou não. O simbolismo dos contos é vivido pela criança como uma resposta aos seus mais profundos medos e anseios. Os contos oferecem soluções simbólicas para os problemas que a criança vai experimentando durante o seu processo de desenvolvimento e de consciência do eu interior. Elas concebem e estruturam no seu pensamento as respostas simbólicas exprimidas de forma abstracta pelos símbolos que lhes são oferecidos pelos contos. Estes, por sua vez, ao oferecerem tais modelos de abstracção, permitem-lhes não apenas viver ou incluir-se numa realidade visível ou, se se preferir, possível, mas também «viver» uma outra realidade invisível onde todos os sonhos se realizam. Utilizando as palavras de Vladimir Propp, «Neste sentido, os contos populares jogam com uma espécie de cumplicidade cultural secundária e encantam na medida em que neles nos identificamos e sobre eles projectamos arquétipos profundos que a narrativa conta sob o pre-texto duma intriga factual, variável. É aqui que o *não-dito* da narrativa assume uma importância fundamental» (Propp, 2003: 24).

Deixemos que as nossas crianças cultivem a sua linguagem simbólica e dela possam fazer livre uso. Disso depende o bom encadeamento de toda uma sucessão de etapas de crescimento que culminará na realização individual e na formação de um ser mais seguro, e não estéril de uma imaginação tão vital como o próprio acto de respirar. Deixemos, portanto, que cada criança possa tornar-se «sonhador de palavras, sonhador de poemas, sonhador de mitos e, assim, instalar-se plenamente na realidade antropológica bem mais vital, mais importante para o destino e, sobretudo, para a *felicidade* do homem do que a morta verdade objectiva. Porque é entre as verdades objectivas desmistificadoras e o insaciável querer ser construtivo do homem que se instala a liberdade poética, a liberdade “remitificante”. Mais do que nunca, nós sentimos que uma ciência sem consciência, isto é, sem afirmação mítica de uma Esperança,

marcaria o declínio definitivo das nossas civilizações» (Durand, 1993: 108 e 109).

5. O estudo informático-lexical dos contos

Os contos da tradição oral passam de geração em geração, imunes a todas as alterações sociais, mantendo vivo um imaginário único que apenas eles conseguem transmitir. Esta magia dos contos vive nas e pelas palavras e temos a convicção de que a oralidade atribui aos contos matizes incomparáveis, que nenhuma escrita poderá jamais reproduzir. Nós, na impossibilidade de obtermos resultados objectivos, rigorosos e sistematizados dos sentimentos e das emoções que eles continuam a transmitir, pretendemos, contudo, obter resultados objectivos, rigorosos e sistematizados em relação ao léxico de que se compõem. Iniciamos assim, aliando tradição e modernidade (ao utilizar as ferramentas da linguística computacional), uma nova forma de estudar os contos da tradição oral.

As vantagens do uso das novas tecnologias no estudo de textos são evidentes. Por um lado, evita toda a interpretação individual baseada em «intuições»; por outro, apresenta resultados sistematizados e rigorosamente fiáveis. Considerando que o acto de leitura, interpretação e análise de qualquer obra resulta da percepção e identificação de índices (linguísticos, paralinguísticos) e do papel das memórias (imediate e mediata), o que permite construir hipóteses de significação a serem verificadas e validadas no decurso da leitura – análise da obra, e considerando também que a leitura, interpretação e análise de qualquer obra depende ainda do estado de espírito de quem lê, da sua «história de vida» e até de um universo de referências e de um repertório linguístico específicos, inferimos que não há uma leitura de um texto, mas várias, tantas quantas, não só o número de leitores, como os momentos de leitura. A aplicação de um programa informático de análise textual permite apagar o mais possível a inevitável filtragem do referido processo de leitura, interpretação e análise de qualquer texto.

O registo e tratamento dos textos sobre suporte informático permitem assegurar um inventário rigoroso e exaustivo do vocabulário do corpus. Graças aos programas de análise estatística, que estabelecem a necessária distância entre os materiais e o investigador, conseguimos resultados objectivos e sistematizados, garantindo uma apresentação neutra e objectiva dos dados quantificados. É ainda o computador que nos permite medir a frequência relativa da palavra, localizar em contexto textual as formas lexicais seleccionadas, caracterizá-las em função de parâmetros de situação e delinear-lhes o perfil numa textura discursiva determinada.

Com as palavras nos construímos e nos damos a conhecer. É também com palavras que revelamos ao mundo o ser que somos. Os contos «revelam-se» também, e

essencialmente, pelas palavras utilizadas e pela frequência com que ocorrem: «a linguagem é um sistema probabilístico, cuja face mais notável é a frequência do uso das palavras [...]. A frequência de uso (alta, baixa, intermediária), atributo inseparável da palavra, pois revela a sua ocorrência observada, tem um papel definidor da palavra, fornecendo um traço tão inseparável quanto o sentido» (Sardinha, 2004: 162).

A língua, ao exercer uma poderosa influência no modo como as pessoas apreendem o mundo exterior e a ele se adaptam, constitui uma das áreas de conhecimento mais importantes dos seres humanos. O vocabulário que dominamos influencia em grande parte a esfera da nossa experiência e a sua estruturação, ao mesmo tempo que a dá a conhecer. Há muito que os críticos literários reconheceram a importância ideacional e temática do vocabulário em textos. A este propósito, João David Pinto Correia atribui a Julia Kristeva e a Roland Barthes o «mérito de “deselitizar” o conceito» de texto ao «adoptarem a descrição do mesmo como “prática significante”» (Correia, 1992: 103). Contudo, esta tarefa não se afigurou fácil, já que, como nos diz Roland Barthes, «definir literatura como “linguagem” seria ofender o seu valor humano [...], negar ou diminuir ao mesmo tempo o seu poder realista [...] e o seu poder poético» (Barthes, 1968: 10). Actualmente, a aproximação entre a literatura e a linguística não só é considerada natural como imprescindível: «El comentario lingüístico supone analizar el funcionamiento de la lengua en el texto dado; posteriormente se intenta llegar a una síntesis, cuya fórmula general sería la siguiente: el texto en cuestión dice lo que dice en virtud de la particular realización del sistema lingüístico que en él se observa» (Alconcel, 1993: 19). Por outro lado, ou melhor, noutra aceção, nas palavras de Roger Fowler, «se pensarmos no vocabulário ou no léxico como uma codificação de ideias ou experiências, ele adquire uma qualidade muito mais dinâmica e deixa de ser uma mera lista de palavras inactivas num dicionário» (Fowler, 1994: 259).

Os contos da tradição oral já foram objecto de estudo de disciplinas como a História, a Etnologia, o Folclore, a Literatura, a Linguística, a Psicologia e a Psicoterapia. E, ainda que, já no âmbito da Linguística, tenham sido alvo de alguns estudos (como o aprofundado estudo de Vladimir Propp⁵ sobre a estrutura do conto, em *A Morfologia do Conto*, que, de resto, se tem mantido como uma obra de referência para quantos trabalham os contos), ainda não foram objecto de um estudo informático-lexical e simbólico, com o objectivo de aferir do léxico que, segundo acreditamos, é responsável pela importância que têm no desenvolvimento da criança e pelo fascínio que continuam a exercer no ser adulto.

⁵ Seguido por dois estudos dentro da mesma linha de estudo morfológico. O primeiro de Alan Dundes (que aprofunda o estudo de Propp, revelando combinações mais complexas a partir de 6 motivos fundamentais) e o segundo dos franceses Bremond e Greimas (em que Greimas iniciará um estudo baseado na análise semântica dos contos populares).

Utilizando as ferramentas de que dispomos, e com a ajuda acrescida de métodos estatísticos, é possível criarmos e interpretarmos dados quantitativos, cuja análise nos permite um conhecimento aprofundado, não apenas da forma, mas da substância do corpus que estudamos, e que não seria possível efectuar (ou obteríamos dados não confiáveis) de outra forma.

Graeme Kennedy, em *An Introduction to Corpus Linguistics*, reconhece o contributo das análises manuais de textos ao longo dos séculos, especialmente para a lexicografia, mas fala claramente na vantagem do uso do computador e da fiabilidade dos resultados obtidos neste segundo método sobre o primeiro: «The analysis of huge bodies of text 'by hand' can be prone to error and is not always exhaustive or easily replicable [...]. The Corpus Linguistics is thus now inextricably linked to the computer, which has introduced incredible speed, total accountability, accurate replicability, statistical reliability and the ability to handle huge amounts of data» (Kennedy, 1998: 5).

Com estas ferramentas poderemos verificar, com toda a fiabilidade, a frequência com que determinadas palavras ocorrem no contexto dos contos, verificar quais as palavras-chave, as palavras-tema, os termos exclusivos ou formas de frequência 1 (os *hapax legomena* ou *hapaxes*⁶), etc. Com dados relativos a cada autor poderemos comparar as ocorrências lexicais entre os diversos textos, a riqueza de vocabulário existente num e nos outros casos e estabelecer concordâncias com toda a fiabilidade. Segundo Tony Berber Sardinha, “o campo do léxico [...] possui a maior quantidade de conhecimento derivado do exame de corpora” (Sardinha 2004: 24).

Abordar os contos da tradição oral sob uma perspectiva linguística (e, no nosso caso, utilizando as ferramentas da linguística computacional), permite-nos tanto o estudo da língua como o estudo linguístico da língua. Com efeito, e nas palavras de M.A.K. Halliday:

Ao falarmos, portanto, da abordagem linguística de textos literários, não nos referimos meramente ao estudo da língua, mas sim ao estudo desses textos segundo os métodos da linguística. Há uma diferença entre fazer afirmações *ad hoc* de tipo textual ou linguístico acerca da literatura, que assumem um carácter pessoal e seleccionam elementos de forma arbitrária – como as que podem aparecer a apoiar a tese literária pré-estabelecida – e descrever um texto com base numa teoria linguística geral (Halliday, 1967: 217-223).

Um corpus, por si só, não nos fornece qualquer informação, é apenas «a store of used language» (Hunston, 2002: 3). No entanto, o software disponível para o tratamento de corpora dá-nos a possibilidade de «re-arrange that store so that observations of various kinds can be made [...]». A corpus does not contain new

⁶ «the words that occur only once are called hapax legomena or hapaxes» (Hunston, 2002: 17) e constituem, segundo Tony Berber Sardinha, a maioria em qualquer corpus (Sardinha, 2004: 24).

information about language, but the software offers us a new perspective on the familiar» (ibid.: 3).

O corpus que estudamos é, segundo as tipologias estabelecidas, um corpus a que Susan Hunston chamou «A corpus of texts of a particular type» e que, por isso, denomina de «Specialised corpus», por comparação com os restantes tipos de corpora que estabelece na sua classificação de tipos de corpora: «Comparable corpora, Parallel corpora, Learner corpus, Pedagogic corpus, Historical or diachronic corpus e Monitor corpus» (Hunston, 2002: 14-16). No mesmo sentido temos, também, a classificação estabelecida por Graeme Kennedy, que define «specialized corpora» como «Corpora which are designed with particular research projects in mind». Segundo Kennedy, os corpora, para além dos já citados «specialized corpora», podem ainda assumir variadíssimas outras designações, dependendo dos critérios segundo os quais são classificados. Neste autor podemos encontrar associadas aos corpora electrónicos expressões como «general corpora, training corpora, test corpora, dialect corpora, regional corpora, non-standard corpora, learner's corpora, written corpora, spoken corpora, sample-test corpus, full-text corpus, dynamic corpora, monitor corpora, synchronic corpora e diachronic corpora» (Kennedy, 1998: 19-22). Já Tony Berber Sardinha, um dos mais recentes autores a estabelecer uma tipologia de corpus, embora vá no sentido da classificação estabelecida pelos anteriores, fá-lo em termos mais complexos, agrupando-os por critérios de classificação. Segundo este autor, a tipologia de corpus seria estabelecida em termos dos seguintes critérios: «modo (falado ou escrito), tempo (sincrónico, diacrónico, contemporâneo, histórico), selecção (de amostragem, monitor, dinâmico, estático, equilibrado), conteúdo (especializado, regional ou dialectal, multilingue), autoria (de aprendiz, de língua nativa), disposição interna (paralelo, alinhado) e finalidade (de estudo, de referência, de treinamento ou teste)». Podemos, segundo o autor, estabelecer esta classificação por meio das respostas às perguntas relacionadas com: «Pluralidade de autoria, origem da autoria, meio, especificidade, dialecto, equilíbrio, fechamento, renovação, temporalidade, tradução e intercalação» (Sardinha, 2004: 20-22). O corpus que estudamos é, também segundo a classificação deste autor, um corpus especializado, tanto no que respeita ao conteúdo, como nas possíveis respostas à pergunta relativa à especificidade. O corpus de que nos ocupamos é, de facto, composto por um tipo específico de textos, no caso, textos de contos da tradição oral. Utilizando uma nomenclatura mais completa em termos de definição do corpus de que nos ocupamos, podemos, para além de *especializado*, denominar o conjunto dos nossos textos como um corpus escrito (composto de textos escritos), diacrónico (compreende vários períodos de tempo), de aprendiz (autores não nativos), paralelo (textos comparáveis) e de estudo (quanto à sua finalidade).

Trataremos, portanto, do estudo das palavras utilizadas nos contos, individualmente ou no seu conjunto, enquanto vocabulário utilizado pelos autores, e inseridas num contexto mais vasto que é o léxico possível de ser utilizado.

6. Fragmentos dos resultados obtidos pela aplicação dos programas

Um total de 319 textos da tradição oral foi tratado pelo *Lexicon* e pelo *NooJ*, tendo desta forma sido possível aceder ao léxico presente nos contos da tradição oral, sob uma perspectiva objectiva e fiável.

O corpus que estudámos é constituído por 196970 ocorrências, 12473 formas diferentes e 4526 lemas, distribuídos, segundo a sua frequência, por verbos (57883 ocorrências), nomes comuns (39560 ocorrências), conjunções (28166 ocorrências), artigos (27710 ocorrências), preposições (20648 ocorrências), advérbios (13792 ocorrências), adjectivos (4647 ocorrências), numerais (1480 ocorrências), nomes próprios (1428 ocorrências), pronomes (1053 ocorrências), interjeições (413 ocorrências), onomatopeias (150 ocorrências), palavras latinas (36 ocorrências) e castelhanismos (4 ocorrências):

Classe gramatical	TOTAL DOS CONTOS					
	Ocorrências	Média	Formas diferentes	Média	Lemas	Média
Artigos	27710	14,068%	8	0,064%	2	0,044%
Preposições	20648	10,483%	46	0,369%	22	0,486%
Pronomes	1053	0,535%	168	1,347%	46	1,016%
Conjunções	28166	14,300%	19	0,152%	19	0,420%
Numerais	1480	0,751%	37	0,297%	28	0,619%
Interjeições	413	0,210%	14	0,112%	14	0,310%
Advérbios	13792	7,002%	154	1,235%	146	3,226%
Nomes comuns	39560	20,084%	3353	26,882%	2348	51,878%
Nomes próprios ⁷	1428	0,725%	392	3,143%	82	1,812%
Adjectivos	4647	2,360%	686	5,500%	441	9,744%
Verbos	57883	29,387%	7558	60,595%	1349	29,806%
Onomatopeias	150	0,076%	30	0,241%	25	0,552%
Palavras latinas	36	0,018%	4	0,032%	4	0,088%
Castelhanismos	4	0,002%	4	0,032%	- ⁸	-
Totais	196970	100%	12473	100%	4526	100%

Tabela 1 – Totais do Corpus: Ocorrências, Formas diferentes e Lemas

⁷ Estão contabilizados neste campo todos os títulos dos contos.

⁸ As palavras latinas e os castelhanismos foram incluídos nos lemas em português (desde que existentes no corpus) e foram incorporados num só registo, sem distinção da classe gramatical a que pertencem.

Como podemos verificar pelos dados obtidos, na coluna das ocorrências, que corresponde a todas as palavras existentes no corpus, são os verbos que têm o valor mais elevado (57883), seguidos pelos nomes comuns (39560) e pelas conjunções (28166).

Na coluna relativa às formas diferentes, os verbos continuam a ser a classe com frequência mais elevada (7558), seguida pelos nomes comuns (3353) e pelos adjetivos (686).

Na coluna dos lemas, verificamos serem os nomes comuns os mais frequentes (2348), de seguida os verbos (1349) e os adjetivos (441). Neste caso, a menor frequência dos verbos em relação aos nomes comuns deve-se à flexão verbal que, na lematização, se vê reduzida a uma só forma para cada verbo, o infinitivo impessoal.

A enorme diferença entre a frequência das ocorrências das palavras gramaticais e a frequência das formas diferentes e dos lemas das mesmas deve-se ao facto de estas, na língua portuguesa, serem de número reduzido.

O total de formas diferentes (12473) e o total de lemas (4526) constituem, como podemos verificar no gráfico seguinte, uma pequena parte do total de ocorrências do corpus (196790).

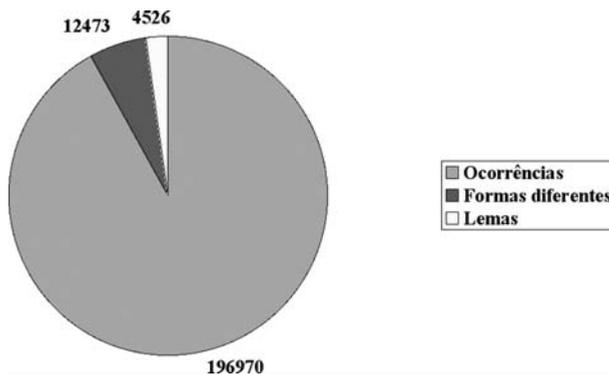
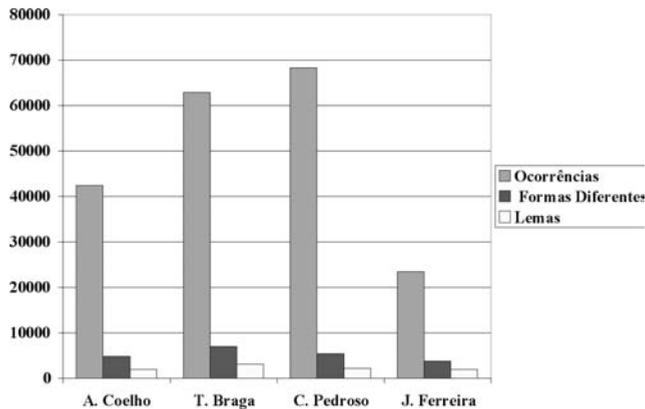


Gráfico 1 – Totais do Corpus: Ocorrências, Formas diferentes e Lemas

Podemos comparar, mais facilmente, os valores relativos ao corpus, ao nível do total de ocorrências respeitantes a cada classe (onde é visível o destaque que os verbos recebem), no gráfico seguinte:

Gráfico 2 – Classes de palavras e categorias gramaticais

Os resultados relativos à contribuição de cada um dos autores para o corpus são os que se apresentam de seguida:

**Gráfico 3 – Totais gerais em cada autor: Ocorrências, Formas diferentes e Lemas**

Riqueza vocabular

Ao nível comparativo, o autor que apresenta maior riqueza vocabular, aferida pela relação percentual entre os totais parciais e os totais gerais, é Joaquim Ferreira (situação a que não deve ser alheio o facto de esta recolha ter sido efectuada quase um século passado sobre as primeiras). A única avaliação que deu resultados diferentes foi a relativa ao uso de termos exclusivos. Nesta, ainda que por uma diferença que considerámos mínima, o autor com maior riqueza vocabular é Teófilo Braga.

Nomes comuns, adjectivos e verbos

Dentro das classes de palavras a que demos maior relevo, os nomes comuns, os adjectivos e os verbos, as formas de maior frequência, aferidas pelas 6 primeiras posições em cada autor, são: *rei, menina, princesa, filha, mulher, rapaz, pai, príncipe, dia, casa* e *palácio* (relativamente aos nomes comuns), *só, branca, grande, peludo, verdes, linda, contente, triste, forte, rico, bonita, seguinte* e *pequeno* (relativamente aos adjectivos) e *foi, disse, tinha, é, era, e estava* (relativamente aos verbos).

Ainda relativamente a estas classes de palavras, a maioria dos nomes comuns e dos adjectivos encontra-se no singular e, estes últimos, predominantemente, nos graus normal e superlativo absoluto analítico, com resultados aproximados no que às posições posposta e anteposta concerne. Relativamente aos verbos, existe um claro predomínio do pretérito perfeito sobre os restantes modos e tempos verbais. No que diz respeito ao número e pessoa, a terceira pessoa do singular apresenta uma frequência consideravelmente superior aos restantes.

Palavras-tema e campos temáticos

No corpus existem 64 palavras-tema que incluímos nos seguintes campos temáticos: a vida temporal, a família, a moralidade e a realeza. Estes campos temáticos podem simbolicamente abranger tantos campos quantos a imaginação individual permitir. As palavras-tema, num total de 64, têm a seguinte distribuição: 34 formas verbais (18 lemas), 2 adjectivos e 28 nomes comuns.

Como pertencendo ao campo temático da «vida temporal» classificámos as seguintes formas: *homem, menina, mulher, rapaz, velha, tempo, fim, terra, dia, noite, água, caminho, só e vez*; como pertencendo ao campo temático da «família» classificámos as seguintes formas: *filha, filho, mãe, pai, casa e porta*; como pertencendo ao campo temático da «moralidade» classificámos as seguintes formas: *bem e mal*; como pertencendo ao campo temático da «realeza» classificámos as seguintes formas: *princesa, príncipe, rainha, rei, palácio, grande, cima e cavalo*.

Simbolicamente, o campo temático da vida temporal pode abranger conceitos como fragilidade, experiência, sabedoria, percurso, etc.; o campo temático da família pode abranger conceitos como protecção, carinho, apoio, amor, etc.; o campo temático da moralidade pode abranger conceitos como perigo, maldade, virtualidade, escolha, etc.; e o campo temático da realeza pode abranger conceitos como poder, sonho, desejo, magnificência, etc.

Utilizando algumas das palavras-tema (que fazem, afinal, o sentido dos contos), podemos construir um pequeno texto, onde o seu encadeamento nos transmite a ideia que parece transparecer nos contos: O ser humano tem um «tempo» na terra, com um

«caminho», constituído por «dias» e por «noites», pelo «bem» e pelo «mal», que por vezes percorre «só» e cujo «fim» representa uma «porta» de passagem.

O uso do diminutivo

O uso do diminutivo, frequente nestes textos, encontra o seu expoente máximo na forma *velhinha*, a mais frequente das comuns nos autores.

A *velhinha* desempenha nos contos, se não sempre, maioritariamente, uma função de protectora e de encaminhadora na resolução de problemas. Esta velhinha dos contos, à semelhança de todas as velhinhas, tem sempre algo de fundamental «para comunicar», manifesto na frequência de ocorrência dos verbos *dizer* e *perguntar* junto da forma em estudo, ou para fazer, de forma a que a ou as personagens em perigo ou com dúvidas, ou ela própria, consigam resolver o conflito em que se encontram.

O papel dos animais nos contos

Um ⁹	Nos contos, a referência aos animais é muito frequente e portadora de grande	Jardim	119
Rei	1271	Gigante	112
Dia	849	Peixe	109
Príncipe	784	Mundo	107
Casa	742	Torre	105
Velhice	714	Sete	104
Princesa	662	Deus	98
Pai	605	Mar	98
Homem	559	Pássaro (ave)	97
Palácio	549	Burro	96
Mãe	473	Diabo	96
Três	457	Palavra	94
Monte	414	Cantar (canto)	93
Dois	375	Boi	92
Noite	367	Gato	92
Porta	355	Cabelo	91
Cavalo	321	Fada	91
Água	288	Flor	90
Vestido	288	Prometer (Prometeu)	90
Olhar	278	Olho	88
Terra	243	Espada	86

⁹ O símbolo «Um» tem uma ocorrência de 5276, devido ao facto de estarem aqui contabilizadas as ocorrências desta palavra quer enquanto numeral, quer enquanto artigo indefinido.

SÍMBOLO	OCORRÊNCIA	SÍMBOLO	OCORRÊNCIA
Mão	240	Cabra, Cabrito	85
Pé (passo)	240	Pena	83
Tempo	235	Cão	81
Pobreza	201	Pão	81
Preto	190	Galinha	80
Ouro	189	Boca	78
Cabeça	188	Sangue	78
Pomba	164	Subir	78
Cama	144	Caçar	77
Pedra	143	Fonte	73
Leão	142	Cavaleiro	72
Casamento	139	Pesca	69
Janela	136	Vento	66
Amor	133	Poço	62
Mesa (tábua)	130	Cidade	60
Branco	125	Luz	60
Raposa	124	Manhã	60
Anel	121		

significado simbólico, tendo a forma *cavalo* recebido, sob as várias análises efectuadas, um destaque superior às restantes.

As ocorrências comuns aos 4 autores são as seguintes: *aves, boi, bois, burra, burro, cabra, cadelinha, cavalo, cavalos, cobra, coelhinho, coelho, coelhos, formiga, galinha, galinhas, galo, gato, leão, lobo, macho, ovelhas, passarinho, passarinhos, pássaros, patos, peixe, pomba, porca, porco, raposa.*

O léxico simbólico

Sob o ponto de vista simbólico, os textos são de uma enorme riqueza. Os símbolos encontrados com ocorrência superior a 60 são:

LÉXICO SÍMBÓLICO COM OCORRÊNCIA IGUAL OU SUPERIOR A 60			
SÍMBOLO	OCORRÊNCIA	SÍMBOLO	OCORRÊNCIA

Palavras não dicionarizadas

No corpus encontramos 93 palavras não dicionarizadas, com as quais construímos um pequeno «glossário» que consideramos um modesto contributo no âmbito da lexicografia.

Bibliografia

- ALCONCHEL, José Luís Girón (1993). *Introducción a la explicación lingüística de textos*. Madrid: Editorial Edinumen.
- BARTHES, Roland, *et al.* (1968). *Linguística e literatura*. Tradução de Isabel Gonçalves e Margarida Barahona. Lisboa: Edições 70.
- BETTELHEIM, Bruno (1995). *A psicanálise dos contos de fadas*. Lisboa: Ed. Bertrand.
- BRAGA, Teófilo (2002). *Contos Tradicionais do Povo Português*. 6ª edição. V. 1 e 2. Lisboa: Publicações Dom Quixote Lda.
- CHEVALIER, Jean e Gheerbrant, Alain (1982). *Dicionário dos Símbolos*. Lisboa: Teorema.
- COELHO, Adolfo (2005). *Contos Populares Portugueses*. 8ª edição. Lisboa: Publicações Dom Quixote.
- CORREIA, João David Pinto (1992). «Para uma teoria do texto da literatura popular tradicional». In Manuel Viegas Guerreiro (coordenação). *Literatura popular portuguesa, teoria da literatura oral/tradicional/popular*. Lisboa: Acarte. Fundação Calouste Gulbenkian, 101-128.
- DURAND, Gilbert (1993). *A Imaginação simbólica*. Lisboa: Edições 70.
- FERREIRA, Joaquim Alves (1999). *Literatura Popular de Trás-os-Montes e Alto Douro – Lendas e Contos Infantis*. Vila Real: Edição conjunta com a CMVR.
- FROMM, Erich (1980). *A linguagem esquecida: uma introdução ao entendimento dos sonhos, contos de fadas e mitos*. Rio de Janeiro: Zahar.
- FOWLER, Roger (1994). *Crítica Linguística*. Tradução de Mª Luísa Falcão e Isabel Mealha. Lisboa: Fundação Calouste Gulbenkian. Serviço de Educação.
- HALLIDAY, M.A.K. (1967). «The Linguistic Study of Literary Texts». In Seymour Chatman e Samuel R. Levin (org.). *Essays on the Language of Literature*. Bóston: Houghton-Mifflin, 217-223.
- HUNSTON, Susan (2002). *Corpora in applied linguistics*. Cambridge: University Press.
- KENNEDY, Graeme (1998). *An introduction to corpus linguistics*. London: Longman.
- PEDROSO, Consiglieri (2000). *Contos Populares Portugueses*. 7ª edição. Lisboa: Vega.
- PLATÃO (2001). *República*. Trad. e notas de Mª Helena da Rocha Pereira. 9ª edição. Lisboa: Fundação Calouste Gulbenkian.
- PROPP, Vladimir (2003). *Morfologia do conto*. Lisboa: Vega.
- SARDINHA, Berber, T. (2004). *Linguística de corpus*. São Paulo: Manole.
- TRAÇA, Maria Emília (1992). *O fio da memória. Do conto popular ao conto para crianças*. Porto: Porto Editora.