



---

## Para uma análise da lisibilidade de textos escritos em português europeu<sup>1</sup>

Ana Martins (Universidade do Porto)

**RESUMO:** O estudo experimental em descrição neste artigo resulta na proposta de linhas orientadoras para a construção de um instrumentário de avaliação da lisibilidade de textos para leitura extensiva em português L2. Os objectivos do estudo são testar e descrever procedimentos de retextualização, através da conjugação de modelos de análise de textos modificados para o inglês e de descrições sobre factores de complexidade linguística em português. Os resultados deste estudo, a saber, o apuramento dos processos de retextualização mais rentáveis, permitem apontar os parâmetros *frequência de lemas*, *processos derivacionais*, *co-referência* e *estruturas de subordinação* como base para a constituição desse instrumentário.

Palavras-chave: leitura; aquisição de L2; lisibilidade; avaliação

### Introdução

As fórmulas de lisibilidade mais divulgadas (para o inglês, o *Flesch Reading Ease Score* (1948) ou o *SMOG Grading Formula* (1969); para o alemão, o *Lix* (1968); para o espanhol, a *Fórmula de Lecturabilidad* de Fernández-Huerta (1959), por exemplo), têm em comum procederem a uma análise quantitativa simplificada, fundada, basicamente, na extensão da frase, em número de palavras, e na extensão da palavra, em número de sílabas, por frase. Actualmente estas medições não se realizam manualmente, mas através de aplicações informáticas<sup>2</sup>.

Os vários estudos<sup>3</sup> que provam a falibilidade destas fórmulas (revelando, por exemplo, que frases com pontuação aproximada nas fórmulas apresentam índices de frequência de vocabulário e complexidade estrutural muito diferenciados; ou, ainda, confrontando palavras polissilábicas, plenamente conhecidas, com monossílabos, deícticos ou preposições, mas de referência inacessível para o leitor) resultam na conclusão de que nenhuma avaliação de lisibilidade pode prescindir, por um lado, da consideração do texto como objecto coeso e, por outro, do índice de frequência dos lemas.

---

<sup>1</sup> O presente artigo decorre do trabalho realizado no âmbito do projecto de pós-doutoramento financiado pela Fundação para a Ciência e a Tecnologia - Ref. SFRH / BPD / 40498 / 2007.

<sup>2</sup> Ver Flesch - Document Readability Calculator <http://flesh.sourceforge.net/> ; SMOG Calculator <http://www.harrymclaughlin.com/SMOG.htm>

<sup>3</sup> Ver Davison, A. & Kantor, R. N. (1982); Bernhardt (1984); Coady (1993); Carver (1994).

Um dos campos de reaplicação das fórmulas de lisibilidade é o da produção e avaliação de textos para leitura extensiva em L2. No caso do inglês L2, tanto editores como órgãos de gestão educacional pautam a regulação dos materiais de leitura em função de instrumentos de avaliação da acessibilidade dos textos, de acordo diferentes níveis de proficiência. Conversamente, os textos, previamente calibrados, servem depois para aferir o nível de proficiência do aprendente. No entanto, as fórmulas de lisibilidade para o inglês, além da eficácia reduzida que vem sendo constatada, não têm aplicação noutras línguas, designadamente nas línguas românicas, desde logo pelo carácter analítico destas.

O trabalho experimental, que a seguir se descreve, de inventariação de procedimentos de reescrita com aplicação na produção de textos modificados para leitura extensiva em português L2, visa, como resultado ulterior, o apuramento de dados de medição da lisibilidade, dos quais é possível fazer derivar critérios fundamentados para avaliação qualitativa e quantitativa da acessibilidade de textos integrados em programas de leitura em português L2, passíveis, alguns deles, de serem deriváveis para português L1.

## **1. Leitura extensiva em L2: textos originais ou textos modificados?**

A actual teorização sobre leitura em L1 e L2 faz eco, recorrentemente, do paradoxo<sup>4</sup>: se se dá como certo, por um lado, que a leitura extensiva de textos originais propuliona o reconhecimento automático de vocabulário de baixa frequência e de estruturas transfrásicas complexas, também não se dá como menos certo ser pernicioso propor textos de leitura extensiva para os quais os aprendentes ainda não ganharam suficiente automatismo de decodificação, por insuficiente competência gráfica e lexical.

Como saldar a diferença de processamento envolvido na transição da leitura de decodificação/compreensão local para a leitura de compreensão global/interpretação; ou, se quisermos, como proceder à passagem da fase do *aprender para ler* para a fase do *ler para aprender*?

### **1.1. Vantagens do texto (bem) modificado**

O balanço entre posições teóricas antagónicas sobre o uso de textos simplificados, bem como o estudo empírico comparativo sobre estruturas linguísticas em amostras de textos simplificados e em amostras de textos originais correspondentes, levado a cabo por Crossley et al. (2007), legitima a assunção de que o resultado de uma modificação tecnicamente bem realizada – e assumidamente alheia ao objectivo de recriar ou reproduzir a plenitude comunicacional do texto original – respeitará as características essenciais da língua-alvo (vocabulário e padrões de recorrência construcionais) em produção discursiva não manipulada. Um texto modificado pode ser, então, uma manifestação linguística em “estado natural”<sup>5</sup>, desde que realizada de modo estratégico e regulado.

O texto (adaptado) permite o acesso facilitado de uma forma de palavra a um significado, mas também opera construções que reflectem as representações mentais complexas de uma forma de palavra, através do desenvolvimento dos *schemata* adstritos a um

---

<sup>4</sup> Para uma revisão do estado da questão ver Grabe (2002).

<sup>5</sup> Pode considerar-se estes textos como “não naturais” apenas na medida em que, a par de comunicar um sentido, têm o objectivo primordial de ensinar a língua.

lema e da activação de redes vocabulares intra e inter campos lexicais. Paralelamente, são visados, de modo insistente, colocações e matizes de registo atinentes a um lema.

Se se assume que o conhecimento lexical deve ser construído na leitura, através de desenvolvimento de estratégias que reforcem a competência de indução de palpites correctos e a capacidade de associação morfológica (Nation e Coady , 1988; Nation, 1990), então vale a pena construir mecanismos de manipulação de texto destinados justamente a potenciar mais garantias de indução correcta, por paráfrases e adução de sequências de expansão/justificação, pontuadas, doseadamente, por palavras cognatas.

## 2. Trabalho experimental de produção de três textos para leitura extensiva em português L2

Na prossecução dos objectivos de apurar operações prototípicas de *supressão*, *substituição* e *reconstrução*, de e entre sequências textuais, e de fornecer instrumentos de avaliação que suportem a previsão do grau de conforto de leitura potenciado por um texto relativamente a um dado nível de proficiência, foi iniciado, em 2008, um trabalho experimental que passou (i) pela produção de três adaptações de livros de viagens contemporâneos de autores portugueses (*No Dorso do Dragão. Aventuras e Desventuras de uma Portuguesa na China*, de Cláudia Ribeiro, 2001<sup>6</sup>, para o nível B1; *A Mais Alta Solidão*, de João Garcia, 2002<sup>7</sup>, para o nível A2; e *Planisfério Pessoal*, de Gonçalo Cadilhe, 2005<sup>8</sup>, para o nível B1); (ii) edição *online* multimédia<sup>9</sup> destes produtos; (iii) construção de lemários anotados<sup>10</sup> (iv) inventários das colocações activadas e (v) testagem da eficácia dos objectos manipulados junto de aprendentes de PL2 dos níveis A2 e B1.

A disponibilização em versão multimédia das retextualizações destina-se a garantir a eliminação de entropias decorrentes de limitações na competência lexical e de falhas no reconhecimento da relação grafia-fonia de uma forma de palavra e, enfim, pretende facilitar e tornar atractiva a leitura, provocando a adesão dos aprendentes.

Assim, a disponibilização de *links* de imagem e vídeos e de *links* para verbetes adaptados, ancorados em formas de palavra seleccionadas, permite que o fluxo de leitura não seja interrompido pela consulta e interpretação dos verbetes de dicionários gerais, diminuindo assim, em grande parte, o esforço de actualização lexical durante o período da leitura e minimizando o risco de desistência.

Por outro lado, a sonorização dos textos, por locutores profissionais<sup>11</sup> visa elidir problemas decorrentes de uma decifração incorrecta, particularmente no caso de palavras homógrafas ou parónimas. Acresce que a prosódia auxilia, como é sabido, a tarefa de

<sup>6</sup> Doravante T1.

<sup>7</sup> Doravante T2.

<sup>8</sup> Doravante T3.

<sup>9</sup> T1: <http://www.prof2000.pt/users/anamartins/nodorsododragao/> (pedir credenciais de acesso para [anamartins@mail.prof2000.pt](mailto:anamartins@mail.prof2000.pt));

T2: <http://users.janssenweb.com/anamartins/altasolidao/> (pedir credenciais de acesso para [anamartins@mail.prof2000.pt](mailto:anamartins@mail.prof2000.pt));

T3: <http://www.prof2000.pt/users/anamartins/planisferio/> ou <http://users.janssenweb.com/anamartins/planisferio/> (pedir credenciais de acesso para [anamartins@mail.prof2000.pt](mailto:anamartins@mail.prof2000.pt));

<sup>10</sup> T1: [http://users.janssenweb.com/anamartins/?page=vocab&text=1&show=full](http://users.janssenweb.com/anamartins/?page=vocab&text=1&show=full;);

T2: [http://users.janssenweb.com/anamartins/?page=vocab&text=2&show=full](http://users.janssenweb.com/anamartins/?page=vocab&text=2&show=full;);

T3: <http://users.janssenweb.com/anamartins/?page=vocab&text=3&show=full>.

<sup>11</sup> As vozes em T1 e T2 são, respectivamente, de Marina de Castro, da Antena 3, e de Miguel Roque, da Antena 2.

compreensão imediata dos enunciados e sequências de enunciados. Pretende-se, portanto, que, chegados à fase de realização dos exercícios de testagem, os textos estejam depurados de factores externos de lisibilidade (ou seja, que não os decorrentes de construções linguísticas, estritamente).

O inventário progressivo de palavras substituintes, elaborado aquando dos processos de reformulação sintáctica, enunciativa e discursiva, determinada pelas manipulações inerentes à actividade de retextualização, permite um controlo da extensão do vocabulário em activação; paralelamente, é a partir deste inventário que se constituem os lemários anotados.

A base de dados criada a partir dos lemários<sup>12</sup> permite um controlo qualitativo do vocabulário activado, dado que faculta a fundamentação para as opções lexicais – e por conseguinte, estruturais – tomadas, quer no decurso da retextualização, quer após o seu termo (para efeitos de revisão).

A navegação na base de dados permite a leitura selectiva e de distribuição, em função das etiquetas seleccionadas. As etiquetas *reffreq* (frequência de referência, com recurso ao AC/DC> frequência da Linguateca <http://www.linguateca.pt/>), *pf* (presença ou ausência do lema no vocabulário do Português Fundamental) e *ref* (frequência no texto – T1, T2 ou T3), permitem cruzar taxas de frequência e assim avaliar a justeza da selecção de um lema em detrimento de outro lema isofuncional ou semântica e discursivamente equivalente. A etiqueta *coref* (co-referência e relações lexicais de sinonímia e antonímia<sup>13</sup>) e *family* (presença de palavras cognatas) suportam, ainda que parcialmente, a avaliação sobre o grau de redundância e coesão do texto. A etiqueta *class* (classe de palavra) impede que lemas coincidentes na sua forma sejam contabilizados como um só lema.

Finalmente, a listagem de colocações<sup>14</sup> tem em vista uma posterior avaliação de frequência que vá além da contabilização dos lemas, cruzando dados da frequência na língua, com o auxílio do já referido serviço da Linguateca, com os dados da frequência no texto.

## 2.1. Princípios subjacentes ao processo de retextualização

Presidiram ao processo de retextualização três princípios dados como consensuais em aquisição de L2 com recurso ao exercício da leitura extensiva: (i) o papel positivo da frequência de ocorrências; (ii) a necessidade de aproximação a um dado estágio da interlíngua<sup>15</sup>; e (iii) a necessidade de regulação de inferências.

Na observação destes pressupostos, atendeu-se a factores da ordem do acesso ao conhecimento enciclopédico, da ordem do controlo lexical, da codificação proposicional e da reconstrução textual.

### 2.1.1. Acesso ao conhecimento enciclopédico

---

<sup>12</sup> Ver imagem-amostra em <http://www.prof2000.pt/users/anamartins/images/VOCLER.png>

<sup>13</sup> Para a marcação de antónimos recorreu-se ao sinal [-].

<sup>14</sup> T2: <http://users.janssenweb.com/anamartins/altasolidao/Colocacoes.html>

<sup>15</sup> Entende-se por *interlíngua* o sistema linguístico criado pela influência da L1 e, cumulativamente, pelo *input* da L2; um sistema reduzido, dinâmico e variável. (Ver resumo da evolução deste conceito operativo em Savielle-Troike (2006), p. 40-43.

O processo de edição *online* visa directamente a necessidade de o texto ser acompanhado de informações de radicação cultural e histórica<sup>16</sup>. Esse processo envolve:

- a informação, verbal, icónica ou videográfica, oferecida em *link* ao longo do texto, ancorada em vocabulário seleccionado;
- a redacção de introdução, onde se apresenta um breve enquadramento dos temas e acções relatadas;
- a segmentação do texto pautada por títulos e subtítulos.

### 2.1.2. Controlo lexical

O limite de lemas é o constrangimento mais severo no processo de retextualização. O texto tem de ser adaptado com base na previsão de quais são as palavras arquivadas no léxico passivo<sup>17</sup> dos aprendentes. Para tal cruzam-se os resultados da análise de *corpora* de pequenos textos escritos de alunos de português L2<sup>18</sup> com o atendimento à faixa de frequência de lemas (através do recurso ao vocabulário do Português Fundamental e da plataforma AC/DC>Frequência da Linguatca).

De um modo geral, a selecção de lemas ou expressões para ancoragem em *link* de glosas ou de imagens/vídeos obedece ao critério do grau de frequência. No entanto, há situações específicas em que, apesar da baixa frequência de um item lexical, não se opta pela explicação em *link*, com o intuito de se poder aferir o grau de eficácia das paráfrases construídas no corpo do texto ou a efectividade da leitura de outras glosas em que o item em causa é visado.

A limitação do número de lemas que se pressupõe pertencerem já ao vocabulário dos aprendentes (entre 1000 e 1300 para o nível B1 e não mais de 700 para o nível A2) convoca a repetição de formas, ainda que com distribuições diferentes.

As técnicas de gestão do léxico no texto têm em conta uma previsão das competências estratégias do aprendente, designadamente, estratégias de memorização (capacidade de gestão da memória de trabalho), de associação ou de compensação. Por exemplo, visando o fim específico da consolidação/aquisição de um item lexical, opera-se a transição de um termo dado como previamente conhecido para um termo co-referente de mais baixa frequência, em contextos equivalentes – termo que se irá repetir desde o momento em que se dá essa transição (veja-se, a título de exemplo, a correspondência «viajante» – «passageiro», em T3).

Uma vez assegurada a representação mental complexa de uma palavra, recorrente, estima-se uma maior tolerância à extensão da frase e às estruturas clivadas. Deste modo, aposta-se numa complexificação progressiva da estrutura de frase à medida que avança a retextualização.

### 2.1.3. Codificação proposicional

Neste âmbito são visadas as seguintes operações fundamentais:

---

<sup>16</sup> «Experimental findings which do not acknowledge the impact of reader background are questionable. Reading research which is exclusively text-based taps only a proportion of the construct which is allegedly investigating» Bernhardt (1984), p. 329.

<sup>17</sup> Faz-se aqui eco da distinção, adoptada por muitos investigadores, entre conhecimento passivo/receptivo e conhecimento activo/produzido de uma palavra. Ver, por exemplo, Meara (1990) e Nation, (2001).

<sup>18</sup> Estes *corpora* estiveram na base do trabalho desenvolvido por Leiria (2001) e integram agora o *Corpus* de Produções Escritas de Aprendentes de PL2: <http://www.uc.pt/fluc/rcpl2/>

- reduzir, por elisão e modificação, o número de frases com mais do que uma oração finita: por norma, opta-se por não mais de uma oração subordinada e não mais de três orações coordenadas por frase;
  - optar por frases activas, com ordem directa;
  - excluir orações finitas nos complementos nominais;
  - optar por sujeito explícito pelo menos uma vez por sequência intermédia;
  - eliminar formas verbais de conjuntivo, condicional e de pretérito mais-que-perfeito;
- algumas demonstrações:

Estruturas supletivas de conjuntivo – T2		Estruturas supletivas de condicional – T2		Estruturas supletivas de perfeito – T2	
<i>Original</i>	<i>Texto modificado</i>	<i>Original</i>	<i>Texto modificado</i>	<i>Original</i>	<i>Texto modificado</i>
« <b>Para que uma expedição desse lucro...</b> »	«Para uma expedição dar lucro...»	«A distância parecia curta – estaríamos nessa altura aí a uns 250 m do cume – mas demoraríamos pelo menos meia hora a fazer o caminho.»	«A distância parecia pequena. Afinal, eram apenas 250 metros. O problema foi que demorámos meia hora para voltar ao cume.»	«Não tinha a certeza absoluta de ter chegado ao cume.» ---- «Tinha chegado ao ponto mais alto do mundo.»	«Não tinha a certeza se já estava no cume ou não.» ---- «Cheguei, finalmente, ao ponto mais alto do mundo!»

Fig. 1 Alterações na codificação proposicional motivadas pela restrição de formas verbais: exemplos

- aumentar o número de enlaces sequenciais por expansão (com aumento do número elos co-referenciais e de palavras cognatas):

<b>T2</b>	<p>«Finalmente, resolvi <b>descer</b>. Levantei-me e comecei a <b>caminhar para baixo</b>.»</p> <hr/> <p>«<b>No princípio, quando começámos a preparar esta expedição</b>, eu sabia que não podia ir.»</p> <hr/> <p>Bom, <b>naquela altura</b> eu nem sabia bem quem ele era. Sabia que o Wielicki já <b>nesse tempo</b> tinha como objectivo...»</p> <hr/> <p>«No dia seguinte voltei a subir ao cume e <b>desci</b> então para o lado certo, para o abrigo onde estavam as minhas coisas. Na <b>descida</b> passei pelos meus amigos...»</p> <hr/> <p>«É lá que estão os <b>picos</b> mais altos do mundo. É o único local da Terra onde <b>as montanhas sobem a mais de 8 mil metros</b>.»</p>
<b>T3</b>	<p>«Quando comecei a minha volta ao mundo, aquilo que mais me <b>preocupava</b> era escrever bem (...) A minha <b>preocupação</b> era escrever bem, semana após semana.»</p> <hr/> <p>« Quero <b>agradecer também</b> a todos os meus companheiros de estrada. (...) A terminar, o meu <b>agradecimento...</b>»</p> <hr/> <p>«E, no entanto, <b>gosto da alegria</b> das pessoas, da limpeza, dos espaços verdes. <b>Este encanto</b> repete-se sempre que passeio por uma cidade espanhola...»</p> <hr/> <p>« Acabo, então, por <b>comprar</b> apenas uma biografia de Fernão de Magalhães. É uma <b>compra</b> simbólica.»</p>

Fig. 2 Adução de sequências de expansão: exemplos

Paralelamente, são introduzidos *estímulos de interesse* na leitura, como sejam a introdução de exclamações e de perguntas retóricas em monólogo interior, assim como a introdução de discurso directo e a interpelação directa ao leitor.

#### 2.1.4. Reconstrução textual

Reconhecem-se duas modalidades de reconstrução textual, dependendo do nível a que se destina a retextualização e do grau de presença de tecnolectos.

Para o nível B1 (T1 e T3) operacionaliza-se uma modificação por reprodução parafrástica de sequências breves do texto original, antecedida da exclusão, numa dada macroestrutura, das sequências que não podem ser retextualizadas, por limitação de vocabulário; basicamente, todas as descrições e digressões são fortemente reduzidas. Segue-se a análise das estruturas remanescentes, consubstanciada na:

- selecção das palavras e fraseologias a substituir: o critério de substituição de vocabulário reside nos índices de frequência que um dado item lexical patenteia;
- selecção das estruturas frásicas a reformular em consequência do constrangimento lexical ;
- aplicação das operações descritas em 2.3.3.

Para o nível A2 (T2) implementa-se uma modificação por recomposição de sequências longas, em que o grau de elisão de tópicos temáticos diminui, aumentando, no entanto, o grau de distanciamento face à textualização original, por acção de reordenação dos factos relatados e de sínteses.

Tome-se, como exemplo, em T2, o relato dos preparativos da escalada, ainda em Katmandu: no texto adaptado, a indicação do peso da carga a transportar para a montanha serve de ilustração à referência da quantidade de tarefas que os dois organizadores da expedição tinham de gerir; cumulativamente, essa indicação é uma via de suprimir a enumeração de objectos integrantes dessa carga. Porém, no texto original a indicação surge apenas doze parágrafos adiante (p.85).

## 2.2. Exemplificação das manipulações envolvidas quando da retextualização

As manipulações envolvidas repartem-se, pois, pelas áreas da semântica temporal e enunciação, semântica lexical, sintaxe e estruturação textual. Apresenta-se abaixo um exemplo do trabalho de retextualização de uma pequena porção de texto original de T1.

#### Excerto original

(pp. 202-203)

(1) A minha primeira fascinação tremente perante as agudas montanhas da China foi a caminho do troço de Badaling da Grande Muralha, setenta quilómetros a Norte de Pequim. (2) Eu chegara havia muito pouco tempo e tratava-se da minha primeira excursão fora da capital. (3) O caminho para a Grande Muralha não se fazia então por auto-estrada, como agora. (4) Saberão quanto perdem aqueles que a utilizam? (5) A velha estrada era campestre, com curvas e contra-curvas, provocando numerosos solavancos. (6) Sentia-se o sofrimento do autocarro a subir. (7) De repente, quando nada

o fazia supor, vi um monumento enorme, um monumento verde, belíssimo, que brotava da Terra com um ímpeto avassalador. O que era? Fiquei muda, como num sobressalto. **(8)** Era a minha primeira montanha chinesa. Uma montanha! Como era possível crescerem daquele modo? **(9)** Como era possível duvidar que era bem ali a Morada dos Imortais? **(10)** Mal refeita do encantamento, foram surgindo novas montanhas, como que desligadas umas das outras. Eram concentração de poderosa energia, energia pura, verde-jade, dirigindo-se ao Céu. **(11)** Significavam. Afirmavam o Sagrado. **(12)** Eu rodava a cabeça dentro do autocarro, tomada de uma alegria enorme.

(...)

**(13)** Alcançámos a pé a Muralha, pelas escadinhas que serpenteavam na encosta e, estoirados, entrámos na categoria de valentes. **(14)** Estávamos rodeados de ameias, a mil metros de altitude. **(15)** Recordo a minha desilusão com o aspecto de novinha em folha que a Muralha ali ostentava. E com a quantidade de turistas que a pisavam, só comparável à que se encontrava na Cidade Proibida. (...) **(16)** Mas, se esquecêssemos o pormenor e nos detivéssemos na imensidão, o fascínio era inevitável. **(17)** Em Badaling e em Mutianyu, a Grande Muralha fazia ondular as suas ameias e torres de vigia sobre um mar verde de montanhas a perder de vista. (RIBEIRO, 2001, p. 202-203).

### Excerto retextualizado

#### *A Grande Muralha*

**(1)** Senti-me pela primeira vez fascinada com as altíssimas montanhas da China, quando ia a caminho da Grande Muralha (setenta quilómetros a Norte de Pequim). **(2)** Eu estava há pouco tempo no país e aquela era a minha primeira viagem fora da capital. **(3)** Naquela altura, o caminho para a Grande Muralha não se fazia pela auto-estrada, como agora. **(4)** Quem vai hoje pela auto-estrada perde o melhor da viagem. **(5)** Nós fomos por uma velha estrada, pelo meio dos campos, cheia de curvas. **(6)** Sentia-se a força que o autocarro fazia para subir. **(7)** De repente, apareceu, do nada, uma montanha enorme, verde, belíssima, que saía da Terra com uma força admirável. Aquilo era uma verdadeira surpresa para mim: fiquei sem conseguir dizer palavra; estava até uma pouco assustada... **(8)** Era a minha primeira montanha chinesa. Uma montanha! Como era possível crescerem daquele modo? **(9)** Como era possível não acreditar que era ali que moravam os deuses? **(10)** E logo apareceram novas montanhas, uma após outra, subindo para o Céu. **(11)** Aquelas montanhas eram a própria imagem do Sagrado. **(12)** Eu virava a cabeça, para um lado e para o outro, dentro do autocarro. Estava tão alegre!

**(13)** Para chegar à Muralha é preciso ir a pé, por umas escadinhas às curvas. **(14)** Chegámos: estávamos agora a mil metros de altitude. **(15)** Lembro-me que fiquei um pouco triste por ver que ali a multidão de turistas era tão grande como a que estava na Cidade Proibida. **(16)** Bom, pensei eu, o melhor era esquecer isso e ficar a olhar longamente a paisagem. O encanto não podia ser maior: **(17)** a Grande Muralha, deitada sobre o mar verde de montanhas sem fim...

Descrevem-se abaixo os principais procedimentos envolvidos.

**(1)** A supressão /substituição de lexema (“fascinação”) implicou a anulação da nominalização com complemento pesado e a opção por desdobramento oracional (oração temporal); por sua vez, esta opção permitiu a explicitação de “a caminho” em “ia a caminho”; a supressão do topónimo é contígua à necessidade de suprimir o lema “troço” (expansível apenas por paráfrase artificial).



(2) A eliminação da forma de mais-que-perfeito exigiu a substituição da forma verbal (“chegar”, verbo de “accomplishment” por “estar”, verbo de estado); procedeu-se à substituição do imperfeito pelo presente na conjugação do verbo “haver” com o significado de “decurso temporal” (combinação mais frequente); a opção por “estar” exige, por seu turno, a presença de locativo; “excursão” é uma palavra de baixa frequência o que dita a sua não inclusão na lista de lemas e a sua substituição por “viagem”.

(3) Procedeu-se à substituição do adverbial temporal “então” por “naquela altura”: “então” é usado restritivamente nesta retextualização como conector conclusivo (uso mais frequente); o paralelismo *tempo da enunciação – tempo anterior* inicia-se à cabeça da frase (“Naquela altura – agora”).

(4) Foram colocados em equivalência os deícticos “hoje” – “agora” na referência de um intervalo de tempo dilatado que cobre o tempo da elocução/enunciação escrita; a supressão do uso modal de futuro conduziu à explicitação da implicatura da pergunta retórica; eliminou-se a dupla subordinação (interrogativa indirecta seguida de relativa); a repetição do nome “autoestrada” justifica-se pelo facto de a palavra não ter ocorrido até esta fase da retextualização.

(5) Recorreu-se à predicação verbal: o sujeito explicitado marca a transição de uma situação de insaturação referencial (vigente nos dois enunciados imediatamente anteriores) para uma situação de determinação de agentes e acções (retoma de sequência narrativa); tratando-se de uma predicação verbal a presença do oblíquo “pelo meio dos campos” é a opção menos artificial para substituir o adjetivo “campestre” (de baixa frequência); optou-se pela supressão simples da oração gerundiva: trata-se de uma informação de pano de fundo textual cujo conteúdo, a ser mantido, exigiria a construção de uma nova predicação, bem como de paráfrase para a palavra “solavanco”, o que, automaticamente, atribuiria a esse segmento um relevo que ele não tem no texto original.

(6) Substituiu-se o nome “sofrimento” por “esforço” (para efeitos de regulação da limitação de lemas), mantendo assim o animismo; suprimiu-se a perífrase elíptica de progressivo (“a subir”) e optou-se por seleccionar uma oração final não finita (com explicitação da acção operada pela preposição “para”).

(7) A série de substituições lexicais resultou na alteração do número de orações: “quando nada o fazia supor” —> “do nada”; “monumento” —> “montanha”; “brotar” —> “sair”; “ímpeto” —> “força”; “avassalador” —> “admirável”; “muda” —> “sem conseguir dizer palavra”; “sobressalto” —> “um pouco assustada”; projectou-se a sucessão das lexias “ficar” (estado resultante) – “estar” (situação estativa), de modo a induzir a ordenação lógico-temporal dos eventos.

(8) Manteve-se o enunciado original.

(9) As transformações dizem basicamente respeito a substituições lexicais.

(10) Ocorreu a substituição de perífrase aspectual de visão prospectiva (“foram surgindo”) pela colocação “uma após outra”; substituiu-se a oração participial (“mal refeita”) por “e logo” na introdução da série iterativa; esta opção acomoda a redução do número de predicacões.

(11) Optou-se por recorrer a sujeito explícito (SD+ N): a realização do SD no deíctico “aquelas” gera efeito de visualização; suprimiram-se os verbos “significar” e “afirmar”, para observar a limitação de lemas, tendo a reformulação sido feita através da predicação “ser a imagem”.

(12) Optou-se pela paráfrase: “rodar” —> “virar para um lado e para o outro”; a inserção de frase exclamativa justifica-se pela necessidade da supressão da estrutura “tomado de um + N” (pouco frequente).

(13) Optou-se por colocações de alta frequência (“ir a pé”; “é preciso”); ocorreu mudança de registo: *descrição de acções* → *sequência instrucional*; a elisão simples de “entrámos na categoria de valentes” foi necessária porque se trata uma alusão a um episódio anterior que também foi suprimido integralmente.

(14) Explicitou-se a acção “chegar”; projectou-se a sequencialização das situações “chegar” – “estar”; a presença do deíctico “agora” reforça a transicionalidade.

(15) A fusão de frases (e elisão de informação) foi determinada pela limitação de lemas.

(16) A opção por monólogo interior (elocução presentificada) constitui uma estratégia de evitação das formas de imperfeito do conjuntivo; as substituições de lexemas observaram diferentes soluções: “pormenor” → “isso”; “deter” → “ficar a olhar”; “fascínio” → “encanto”; “inevitável” → “não podia ser maior”.

(17) Ocorreu supressão de topónimos; a condensação da construção factiva “fazia ondular as suas ameias e torres de vigia sobre um mar verde” em “deitada” permite manter a associação metafórica; substituiu-se “a perder de vista” por “sem fim”.

Confirma-se, pois, que a totalidade das operações efectuadas tem de conduzir à construção de um objectivo comunicacional autónomo em relação ao do texto original.

### 2.3. Testagem

A fase seguinte do trabalho consistiu na verificação de hipóteses sobre os processos de retextualização mais rentáveis, cruzando a teorização sobre Linguística Textual com alguma informação relativa a técnicas de avaliação<sup>19</sup> em Aquisição de L2.

Para a avaliação global do grau de acomodação do aluno à leitura da versão modificada de T1<sup>20</sup>, foi aplicado um teste, disponível *online* na página respectiva, que agrega a compreensão global (perguntas de I a V), o domínio de vocabulário (pergunta VI), a capacidade/possibilidade de operar inferências correctas (pergunta VII) e a detecção de valores pragmáticos (pergunta VIII).

A população-alvo de testagem do material de leitura constituiu-se de alunos do nível B1, distribuídos por três grupos: G1- alunos *Erasmus*, de formação universitária, em contexto de imersão há, pelo menos, um trimestre; G2- alunos de formação universitária fora de contexto de imersão; G3- alunos integrados no sistema de ensino (básico e secundário) português.

Apresentam-se abaixo os resultados da testagem.

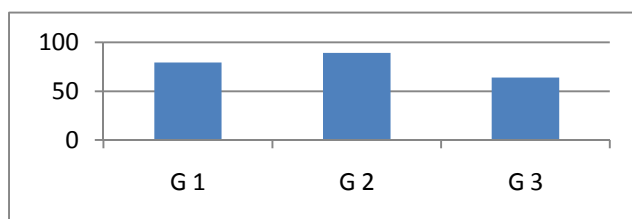


Fig.3 Pontuação global

<sup>19</sup> Laufer (2001); Nation e Laufer (1999); Bachman e Cohen (1998).

<sup>20</sup> T2 e T3 estão ainda em fase de testagem.

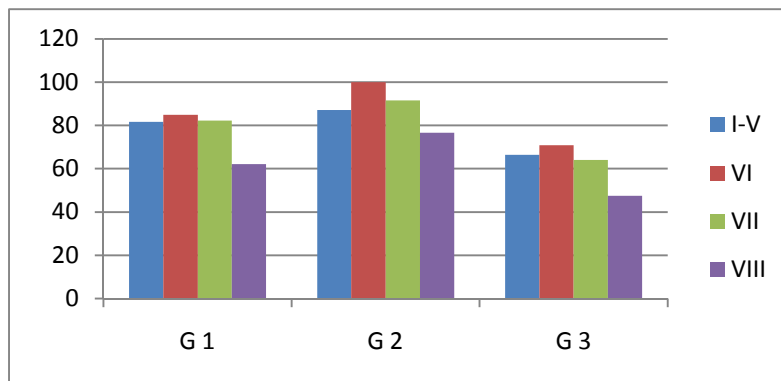


Fig.4 Pontuação parcial

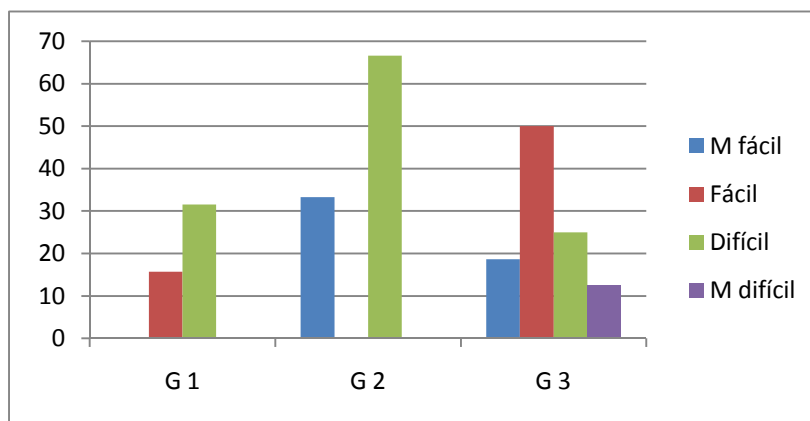


Fig. 5 Avaliação do grau de dificuldade

Apesar da clivagem entre os resultados dos alunos adultos universitários e os dos alunos adolescentes integrados no sistema de escolar (básico e secundário) português<sup>21</sup>, é de assinalar que, nos três grupos, apenas no parâmetro VIII o G3 tenha descido abaixo dos 50%, em 2,5 pontos percentuais, e que todos os demais resultados se tenham mantido acima dos 62%.

Por outro lado, verifica-se que o resultado relativo à avaliação do grau de dificuldade pelo aluno é despicienda, para efeitos de avaliação da lisibilidade, pois não acompanha os resultados obtidos no teste. No entanto, esta avaliação permite aduzir informação acerca da motivação e atitude face às tarefas propostas, informação que pode ser tida em conta na apreciação de resultados de conjunto (integrando os resultados de testagem de T2 e T3).

### 3. Alinhamento dos parâmetros gerais envolvidos na avaliação de lisibilidade

Os resultados deste estudo experimental permitem, de modo consistente, apontar os parâmetros *frequência de lemas (na língua e no texto)*, *processos derivacionais*, *co-referência* e *estruturas de subordinação* como base para a constituição de grelhas que sustentem a validação de previsões acerca do grau de conforto de leitura que um dado texto oferece a um aprendente de um dado nível de proficiência. São estes, justamente, os parâmetros visados na anotação dos *corpora/lemários* resultantes das retextualizações. Dado que a base de dados

<sup>21</sup> Esta clivagem era previsível. Clarke (1980) mostra que leitores proficientes em L1 são-no também em L2 e que os não proficientes em L1 também o não são em L2.

permite a leitura selectiva e de distribuição, em função das etiquetas seleccionadas, esta afigura-se como uma ferramenta útil numa primeira avaliação do grau de lisibilidade de outros textos, modificados ou não.

O refinamento de um instrumentário desta ordem deve pressupor a avaliação da densidade ideal de palavras novas, para o aprendente, presentes no texto<sup>22</sup> e ponderar o peso relativo dos demais parâmetros (derivação, co-referência e subordinação).

Ainda que apenas permita obter resultados por aproximação, dadas as propriedades textuais não tangíveis de parâmetros como as propriedades relativas ao estilo, registo e inferência, um instrumento deste tipo sustentará avaliações mais objectivas de manuais escolares e de materiais suporte no âmbito da investigação sobre aprendizagem da leitura e processamento da informação.

ABSTRACT: The experimental study presented in this paper contributes to set up the basic guidelines towards the construction of a complex tool to measure readability of texts for extensive reading in Portuguese L2. This study aims to test and describe rewriting procedures, by combining frames of analysis of graded readers in English with descriptions about linguistic complexity in Portuguese. The results of this study, namely the elicitation of the most effective rewriting procedures, show that *word frequency, derivation, co-reference* and *embedded structures* should be considered in the process of building such a tool.

Key-words: reading, L2 acquisition, readability, measurement

## Referências

BACHMAN, L. F.; COHEN, A. D. *Interfaces between Second Language Acquisition and Language Testing Research*, Cambridge, Cambridge University Press, 1998.202p.

BERNHARDT, E. B. Towards an Information Processing Perspective in Foreign Language Reading. *The Modern Language Journal*, v. 68, n.4, p. 322-331, 1984.

CARVER, R. P. Percentage of unknown vocabulary words in a text as a function of the relative difficulty of the text: implications for instruction. *Journal of Reading Behaviour*, v.26, p. 413-437, 1994.

CLARIDGE, G. Simplification in graded readers: Measuring the authenticity of graded texts. *Reading in a Foreign Language*, v. 17, n-2, p. 143-158, 2005.

CLARKE, M. A. The short circuit hypothesis of ESL reading. *Modern Language Journal*, v. 64, p.203-09, 1980.

COADY, J. Research on ESL/EFL vocabulary acquisition: putting in context. In.: HUCKIN, T. ; HAYNES, M.;COADY, J. (Eds.). *Second Language reading and vocabulary learning*, Norwood NJ, Ablex Publishing Corporation, 1993. p.3-23.

CROSSLEY, S. et al. A Linguistic analysis of simplified and authentic texts, *The Modern Language Journal*, v. 91, n.1, p.15-30, 2007.

DAVISON, A. & KANTOR, R. N. On the failure of readability formulas to define readable texts: a case study from adaptations. *Reading Research Quarterly*, v.17, n.2, p.187-209, 1982.

---

<sup>22</sup> Hirsh e Nation (1992) e Nation (2001:150): os textos de leitura extensiva não devem conter mais de 5% de ocorrências desconhecidas e não menos de 2% (excluindo os nomes próprios).

FLESCHE, R. A new readability yardstick. *Journal of Applied Psychology*, v.32, n. 3, p.221-233, 1948.

GRABE, W. Dilemmas for the Development of Second Language Reading Abilities. In.: RICHARDS, J. C.; RENANDYA, W. A. (Eds.). *Methodology in Language Teaching. An Anthology of Current Practice*, Cambridge, Cambridge University Press, 2002. p. 276 -285.

HIRSH, D.; NATION, P. What vocabulary size is needed to read unsimplified texts for pleasure?, *Reading in a Foreign Language*, v.8, p. 689-696, 1992.

LAUFER, B. Quantitative evaluation of vocabulary: how it can be done and what it is good for. In.: ELDER, C. et al. (Eds.). *Experimenting with Uncertainty. Essays in Honour of Alan Davies*, Cambridge, Cambridge University Press, 2001. p. 241-250.

LEIRIA, I. *Léxico: aquisição e ensino do português europeu língua não materna*, Dissertação (Doutoramento em Linguística) - Faculdade de Letras da Universidade de Lisboa, Lisboa, 2001.

MEARA, P. A note on passive vocabulary. *Second Language Research*, v.6, p.150-154, 1990.

NASCIMENTO, M. F. B.; MARQUES, M.L.G; CRUZ, M. L. S. *Português Fundamental: Vocabulário e Gramática*. Centro de Linguística da Universidade de Lisboa, 1984.

NATION, P. *Teaching and Learning Vocabulary*, Rowley, MA: Newbury House, 1990.275p.

----- *Learning Vocabulary in Another Language*, Cambridge, Cambridge University Press, 2001.477p.

NATION, P.; COADY, J. Vocabulary and reading. In.: CARTER, R.; McCARTHY, M. (Eds.) *Vocabulary and Language Teaching*, New York, Longman, 1988. p. 97-110.

NATION, P.; LAUFER, B. A vocabulary test of controlled productive ability. *Language Testing*, v. 16, n.1, p. 33-51, 1999.

SAVILLE-TROIKE, M. *Introducing Second Language Acquisition*, Cambridge, Cambridge University Press, 2006.207p.

### **Obras retextualizadas**

CADILHE, Gonçalo 2005 Planisfério Pessoal, Lisboa, Oficina do Livro.

GARCIA, João 2002 A Mais Alta Solidão. O primeiro português do cume do Evereste, Lisboa, Publicações Dom Quixote.

RIBEIRO, Cláudia 2001 No Dorso do Dragão. Aventuras e Desventuras de uma Portuguesa na China, Lisboa, Europa - América.

RECEBIDO EM 31/10/2010 – APROVADO EM 13/05/2011