

## OS RECURSOS E A LINGUAGEM DA WEB

Bento Carlos Dias-da-Silva\*  
Delfina Cristina Paizan\*\*

**RESUMO:** *A Internet e seus recursos vêm sendo explorados por diversos setores da sociedade global da informação. Este trabalho visa a desvendar, sobretudo para o especialista em Letras, uma pequena parte do que subjaz à essa teia virtual de informações, fascinante para alguns, perniciosa para outros e envolvente para os que investigam seus meandros. Inicialmente, apresentam-se um panorama da Internet e sua face gráfica: a Web. Em seguida, aborda-se o potencial de pesquisa dessa megaestrutura de informação e linguagens. Por fim, apresentam-se conclusões e recomendações que devem nortear a publicação de conteúdos em hipertexto, sua forma de expressão.*

**PALAVRAS-CHAVE:** *Internet; Web; hipertexto; HTML; linguagem.*

É cada vez mais comum atribuir-se à rede mundial de computadores, a Internet, a qualidade de ser ela um dos fenômenos tecnológicos mais espetaculares da história. As cifras, indiscutivelmente, indicam que se trata de uma invenção revolucionária, que vem modificando definitivamente o modo de produção, armazenamento e circulação da informação, sobretudo aquela veiculada pelas línguas naturais. Embora não seja tarefa trivial, a extração de estatísticas para aferir os números da Internet aponta cifras da ordem de milhões e bilhões. Diariamente, mais de 300 milhões de

\* Universidade Estadual Paulista - Unesp.

\*\* Universidade Estadual do Oeste do Paraná - Unioeste.

pessoas conectam-se à Internet e nela consultam mais de 3 bilhões de documentos. Estatísticas sobre o uso da Internet e de seus recursos para o ensino-aprendizagem de línguas estrangeiras, sobretudo o inglês, reforçam esse dados e mostram que há mais de 1 bilhão de páginas eletrônicas armazenadas em milhares de servidores (*computer servers*) espalhados pelo mundo, 80% das quais publicadas em inglês (Luft, 2000). Números que confirmam a globalização de significativa parcela da informação que circula no planeta.

Não se pode negar que a Internet permite que as mais variadas atividades como, por exemplo, o acesso a bases de dados e de textos e a interação entre os seus usuários, sejam realizadas de qualquer lugar do mundo por meio de apenas um microcomputador a ela conectado (Warschauer, 1996). Já seus diferentes recursos – o *e-mail*, a transferência de arquivos, a publicação digital, entre outros, vêm sendo usados paulatinamente nas mais diferentes áreas do conhecimento e setores de produção. Embora não tenham sido projetados para aplicação direta no processo de ensino-aprendizagem, esses e outros recursos vêm sendo cada vez mais empregados para esse fim.

## **1. A Internet e suas potencialidades**

A *Internet Society* (2000a, 2000b) apresenta a Internet como uma “rede global de redes que capacita microcomputadores a se comunicar e compartilhar serviços por todo o mundo”, constituindo “um recurso global para o compartilhamento de informação e conhecimentos, assim como de meios de colaboração e cooperação entre comunidades diversas”.

### **1.1 Uma rede indestrutível**

A concepção inicial de projetar uma rede de microcomputadores conectados entre si e de possibilitar, com essa configuração,

aos usuários desse novo sistema acesso a dados e programas a partir de qualquer localidade foi apresentada, em 1962, por J. C. R. Licklider (*Internet Society*, 2000), um dos integrantes da DARPA,<sup>1</sup> agência governamental americana que, durante os obscuros tempos da Guerra Fria, conseguia seduzir importantes pesquisadores. Nesse período da história americana, o objetivo era manter *sites* militares eletronicamente conectados entre si em uma rede de computadores, com a particularidade de nenhum deles ser o comando central. Assim, caso ocorresse um ataque inimigo a algum dos pontos da rede, outros se encarregariam de manter as conexões ativas.

A rede *Arpanet*,<sup>2</sup> antecessora da Internet, atendeu a esse objetivo, servindo como ambiente de teste de tecnologias de rede, conectando universidades e centros de pesquisa. Os dois pontos iniciais dessa rede acadêmica foram instalados em duas instituições universitárias do Estado da Califórnia: a Universidade da Califórnia em Los Angeles e a Universidade de Stanford. Mais tarde, outros dois pontos foram somados: na Universidade da Califórnia em Santa Barbara e na Universidade de Utah, localizada em outro Estado americano. Em 1969, mais quatro pontos juntaram-se a essa rede. Sua apresentação ao público ocorreu em 1974.

Somente em 1982 passou a ser conhecida como Internet (PBS, 2000) e, a partir de então, passou a fazer parte da vida diária de universidades e corporações. No Brasil, os primeiros testes foram feitos em 1991 e, no ano seguinte, já havia uma conexão entre a Fapesp<sup>3</sup> e a UFRJ<sup>4</sup> (NIB, 2000). Indiscutivelmente, seu crescimento tem sido vertiginoso, reunindo, hoje, aproximadamente 390 milhões de usuários, dos quais mais de 9 milhões estão no Brasil (*Global Reach*, 2000).

---

<sup>1</sup> *Defense Advanced Research Project Agency*.

<sup>2</sup> *Advanced Research Projects Agency Network*.

<sup>3</sup> Fundação de Apoio à Pesquisa do Estado de São Paulo.

<sup>4</sup> Universidade Federal do Rio de Janeiro.

A relativa facilidade de acesso tem garantido seu crescimento: um microcomputador, equipado com um *modem* ou uma placa de rede para a interligação das máquinas. Ambos capacitam o microcomputador a transmitir e a receber dados. O primeiro foi desenvolvido para aproveitar a estrutura de linhas telefônicas existentes (o *modem* converte informações digitais em analógicas para que possam ser transmitidas via telefone); a segunda, que utiliza cabos de ligação específicos, foi desenvolvida para que as máquinas pudessem se comunicar sempre por meio de sinais digitais, acelerando significativamente a transmissão dos dados. Além desses dispositivos, a Internet dispõe de milhares de servidores de rede (*servers*), máquinas responsáveis pelo provimento de conteúdos e acesso à rede por pacotes especiais de programas que permitem a visualização e manipulação da informação no monitor do computador: os *browsers*, isto é, os programas de navegação como, por exemplo, o Netscape, o Internet Explorer, o Opera, entre outros.

Uma vez conectado, o usuário entra em um ambiente virtual onde pode participar das mais variadas atividades: recuperar e armazenar informação e, sobretudo, estabelecer comunicação com os demais usuários. Esta última faceta da Internet tem possibilitado uma aproximação ímpar entre membros de comunidades específicas: entre pesquisadores de diversas partes do país e do mundo, permitindo a esses agentes científico-culturais desenvolver trabalhos mais colaborativos e acelerar os resultados de suas investigações; entre grandes organizações, que buscam, cada vez mais, disponibilizar suas informações para o grande público; e entre educadores, para os quais a busca e compartilhamento de conhecimentos nas mais variadas áreas do saber são iniciativas essenciais.

A Internet também apresenta problemas, relacionados com a massa de informação que dissemina sem qualquer preocupação com a qualidade do que está sendo veiculado, e com a falta de segurança e privacidade dos dados que nela são lançados. Essas fragilidades comprometem não só as inúmeras transações comerciais que, ape-

sar dos problemas, são feitas através dela, mas também o sigilo de informações científicas. Para minimizar esta última deficiência, outras duas redes foram propostas: a Intranet e a Internet2.

Embora ambas tenham as mesmas características da Internet, a primeira delas se diferencia porque é uma rede de uso exclusivo de uma determinada organização, cujo acesso é restrito aos seus membros ou a pessoas autorizadas. Além dessa, outras vantagens dessa rede são o seu custo mais baixo e a facilidade de montagem, manutenção e gerenciamento.

A segunda iniciativa surgiu oficialmente em 1996, quando 34 universidades americanas reuniram-se para formar o Comitê Geral de Trabalho da Internet2 (RNP, 2000). Essa rede tem como objetivo restringir o compartilhamento de informações a membros da comunidade acadêmica, servir de laboratório de teste de transferência de tecnologias para o setor comercial e desenvolver novas aplicações, como a telemedicina, as bibliotecas digitais, os laboratórios virtuais, entre outras. Atualmente, dela pode participar qualquer instituição, desde que recomendada pelo Comitê Geral e enquadrada em uma das quatro categorias: universidades, empresas parceiras, empresas patrocinadoras e membros afiliados. Gestões estão sendo feitas pela Fapesp para que as universidades brasileiras possam integrar-se a essa nova via digital.

## 1.2 Uma caixa de ferramentas e linguagens

Um traço marcante da Internet é o fato de ela estar em ebulição constante, constituindo-se para seus usuários em uma caixa de ferramentas digitais, parte delas apresentadas a seguir.

O *telnet* e o *ftp* são as duas das mais antigas ferramentas da Internet. O *telnet* permite que um microcomputador se comunique com outro, ambos estando em lugares arbitrariamente distintos. Permite acesso a bibliotecas, jornais, programas públicos e muitos

outros serviços, desde que se tenha o endereço e, se necessário, a senha de acesso.

O *ftp* (*File Transfer Protocol*) é um protocolo, isto é, um conjunto de instruções, para a transferência de arquivos, o que o diferencia do *telnet*. Ambos permitem acesso aos conteúdos armazenados em computadores remotos. A diferença entre *telnet* e *ftp* é que aquele não permite intercâmbio de documentos, funcionalidade que deu origem a este.

O *e-mail*, isto é, o correio eletrônico, surgiu no início dos anos 70, quando a *Arpanet* era privilégio de apenas pesquisadores, principalmente para compartilhar dados entre computadores remotos. O *e-mail* é, indiscutivelmente, a ferramenta mais popular da Internet, permitindo enviar e receber mensagens, salvá-las em arquivos e respondê-las. O formato padrão do endereço *e-mail* segue uma convenção universal: nome@domínio.site, em que nome é o nome do usuário, o símbolo @, abreviação da preposição *at* do inglês, indica o nó em que está o domínio, isto é, a origem do acesso, e *site* é o tipo do *site*, que pode ser *.edu*, um *site* educacional, *.com*, um *site* comercial, *.gov*, um *site* governamental, *.mil*, um *site* militar, *.org* um *site* de uma organização. Com frequência, o *site* pode incluir também a sigla do país de origem como, por exemplo, *br* (Brasil), *ca* (Canadá), exceto se o país de origem for os Estados Unidos. O fato de o *e-mail* ter se tornado tão difundido deve-se principalmente à sua versatilidade: tudo que pode ser digitalizado, arquivos de textos, de programas, fotos, slides, sons, por exemplo, podem ser transmitidos por esse meio.

As listas de discussão, *mailing lists*, são compostas de grupos de usuários da Internet que compartilham interesses comuns. O acesso a essas listas pode ser feito por *e-mail*: o usuário envia sua mensagem para um endereço específico, a partir do qual ela é distribuída para todos os demais assinantes da lista. Participar, ativa ou passivamente, desses grupos é relativamente simples: após a identificação, participa-se lendo, encaminhando ou enviando men-

sagens. Análogo ao recurso das listas de discussão, a rede *Usenet* possibilita uma forma de participação em discussões em grupo (*newsgroups*). Uma vantagem deste recurso sobre o anterior é que as mensagens não ocupam espaço no computador do usuário, mas ficam concentradas no servidor responsável pela rede *Usenet*.

As ferramentas de comunicação mencionadas não exigem que os usuários, para manterem a comunicação, estejam simultaneamente conectados à rede. Quando o contato exigir simultaneidade nas conexões, isto é, contato sincrônico, são necessárias ferramentas apropriadas. Destacam-se os *chats* e a vídeo conferência. Trata-se de ferramentas de comunicação em tempo real: *Real-Time Communication Tools*. A vídeo conferência já é um recurso mais sofisticado e por isso menos acessível. Esse meio exige equipamentos adicionais: microfone, câmera de vídeo e máquinas mais sofisticadas. No Brasil, um grande problema para o uso desse recurso é a deficiência constitutiva e estrutural das redes de comunicação, responsáveis pela lentidão no processo de transmissão de vídeo e de áudio, além das morosas conexões.

## 2. A face gráfica da Internet

A Web, ou WWW (*World Wide Web*), foi desenvolvida para que as informações circulassem de modo mais eficiente e amigável pela Internet. Para aqueles que têm acesso à rede há pouco tempo, a Internet e Web são vistas como termos sinônimos. Mas o fato é que a Web constitui a face gráfica da “antiga” Internet, resultado das tecnologias introduzidas pela plataforma *Windows* e similares.

A característica principal da Web é a possibilidade de conexão de diferentes mídias por meio de *links* (elos digitais). Para acessá-los, bastam dois movimentos: selecioná-los e clicá-los. Os *links* podem ser uma letra, um palavra, uma expressão, uma figura, um ícone ou os elementos em destaque em um texto. A ação de clicar um *link* transporta o usuário para uma outra parte do mesmo do-

cumento, para outro documento, para arquivos multimídia ou ainda para qualquer outro ponto da rede.

No ambiente gráfico da Web é possível “navegar” por textos, imagens e sons, que transportam os usuários para os mais diversos mundos virtuais. Merecem destaque os programas de busca, denominados pelos informatas “motores de busca”, construídos para auxiliar o usuário a pesquisar informações. O *Yahoo*, criado em 1994, foi o primeiro programa de busca desenhado para esse fim. Esse programa armazena uma imensa base de dados dinâmica (pois está em constante atualização), que organiza toda a informação que circula na Internet em catálogos de palavras-chave que, por sua vez, servem para indexar todos os endereços das páginas que integram a Web. Além desse, exemplos de outros programas de busca são o *Google*, o *Miner*, o *Altavista* e o *Cadê*.

Esse ambiente, com sua linguagem de ícones, menus, grafismos, quadros, imagens, entre outros, tornou-se popular por facilitar e agilizar a publicação de conteúdos e o acesso a documentos. Para que isso acontecesse, entretanto, as seguintes tecnologias foram aprimoradas (Cheng e Malaika, 1997: 8-11):

- *Web browsers*: programas conhecidos como navegadores. Eles permitem a localização e leitura de documentos na WWW. Hoje os mais conhecidos são o *Netscape Communicator* e o *Internet Explorer*, que podem exibir textos e gráficos e reproduzir vídeos, filmes e sons;
- *HTML (HyperText Markup Language)*:<sup>5</sup> uma linguagem formal de representação da estrutura do hipertexto, composta de um conjunto de comandos que permitem a publicação de documentos na WWW. Os comandos na linguagem *HTML* definem, entre outros elementos de formatação, a estrutura de títulos, parágrafos e segmentos de texto, a posição de ele-

---

<sup>5</sup> Ressalte-se que este trabalho focaliza exclusivamente a linguagem HTML com o objetivo de apenas ilustrar o processo de codificação em hipertexto de textos tecidos em língua natural. Há, naturalmente, outras linguagens mais elaboradas e eficientes, cuja discussão extrapolaria o escopo deste trabalho.



mentos gráficos e tipos cores e tamanhos de fontes. Seus dois recursos principais são o *hiperlink*, o rol de comandos que permitem a conexão entre documentos, e a *hipermedia*, ou o rol de comandos que regulam a inclusão de imagens, sons e animação a textos, criando documentos de multimídia;

- *HTTP (HyperText Transfer Protocol)*: um protocolo de comunicação entre máquinas que fornece, entre outras funcionalidades, instruções para o acesso a documentos expressos na linguagem *HTML*. Essas instruções interpretam os documentos e os transmitem através da *WWW*;
- *URL (Uniform Resource Locator)* é um localizador único para cada documento, indicando qual serviço da Internet está sendo usado. Ele pode ter, por exemplo, o formato *http*, para documentos em forma de hipertexto, ou *ftp*, para transferência de arquivo.

## 2.1 A linguagem dos *links*

Devido às limitações impostas pelas dimensões do monitor do computador, restrito a, no máximo, uma página de texto, e a sua pouca resolução, o que torna o processo de leitura de caracteres uma tarefa extenuante para o leitor, a solução adotada pelos informatas foi a segmentação da informação e a posterior conexão dos diversos segmentos, sejam eles textos ou imagens. Dessa necessidade nasceu a idéia do hipertexto, que fornece os recursos necessários para a estruturação de informação segmentada.

De acordo com Barnes (1994), a idéia do hipertexto foi proposta por Vannevar Bush, em 1945, como uma possível solução para o problema de como organizar grande massa de informação. Com o crescimento do volume de artigos científicos publicados na época e com a impossibilidade de se ler e reagir a eles, Bush sugeriu novas formas de acesso, leitura e armazenamento de todo esse material, criando novos elementos para a interação com os textos: *links*, trilhas desses *links* e teias (*webs*) de tais trilhas.

Nos anos 60, Theodore Nelson criou o termo *hipertexto* para descrever a leitura e a produção de texto não linear na tela do microcomputador. O pesquisador imaginou um sistema *on-line* em que toda a literatura produzida pudesse ser armazenada de forma digital e em que *links* eletrônicos permitissem a conexão entre suas partes. Mesmo sendo os textos em língua natural ainda a base de transmissão de mensagens, o conceito de hipertexto se expandiu: hoje é possível conectarem-se outras mídias, e a conexão entre hipertexto e multimídia passou a ser chamada de *hipermídia*.

O hipertexto pode ser definido tanto de um ponto de vista técnico quanto funcional. Do ponto de vista técnico, é “um conjunto de nós ligados por conexões”, podendo ser esses nós “palavras, páginas, gráficos ou parte de gráficos, seqüências sonoras, documentos complexos, que, por sua vez possam ser outros hipertextos” (Levy, 1997: 33). Os itens de informação nessas conexões não são ligados linearmente, como uma corda com nós, mas em forma de estrela, em que cada nó pode, por sua vez, conter uma rede inteira. Do ponto de vista funcional, o hipertexto é um tipo de programa, escrito para a organização de conhecimentos ou dados e que serve para a aquisição e comunicação de informações.

Seis são as marcas que, segundo Levy (1997), caracterizam o hipertexto:

- estar em constante metamorfose;
- apresentar nós e conexões que sejam heterogêneos, podendo ser, por exemplo, imagens, sons, palavras e que tenham presentes, em seu processo de formação, pessoas, grupos, artefatos, entre outros;
- organizar-se de forma “fractal”, ou seja, desdobra-se e conserva as mesmas características; qualquer nó ou conexão pode ser composto por toda uma rede;
- ser dependente, tanto na sua composição e recomposição, de um exterior indeterminado;

- permitir uma aproximação topológica: a relação entre os nós se dá por proximidade de significados;
- ser acêntrico, pois permite várias entradas e vários acessos diferentes.

## 2.2 A face didático-pedagógica da Web

Para Duncan (1997), com a crescente necessidade por informação rápida e precisa, a linguagem do hipertexto tornou-se uma ferramenta indispensável para a disponibilização de conteúdos na Web. McIntyre e Wolff (1998) empregaram-na para tornar o ensino da linguagem de programação C para os alunos de Ciência da Computação mais interativo. Ao invés de aulas convencionais, com restrições de tempo, de número de alunos, que impedem que o professor incentive a participação ativa de cada um, os autores produziram um material que permite ao aluno “navegar” pelos textos e atividades, experimentando possibilidades e avaliando resultados. Soma-se a isso a possibilidade de o aluno poder tomar decisões quanto à hora mais conveniente para se dedicar às atividades de leitura, ao local de trabalho e ao modo de estudo, individual ou em grupo.

No ensino de língua estrangeira, Sengupta (1996) descreve uma experiência de criação de uma base de dados em hipertexto para ao ensino de produção de textos. O autor destacou dois cuidados nessa tarefa: a necessidade de construir uma base de dados coerente, com um tema específico; planejamento meticuloso da disposição e conexão entre as partes do texto, para evitar que o leitor se perca nos caminhos traçados no hipertexto.

Wu e Tsui (1997) usaram o hipertexto na criação de uma gramática eletrônica para professores de inglês como segunda língua. Os autores concluem que o hipertexto é o melhor recurso para se disponibilizar uma gramática, graças a sua característica relacional e não-linear. A principal desvantagem apontada nesse experimento



é praticamente a mesma citada acima: desnortear o leitor. Por outro lado, os autores enfatizam a facilidade de revisão e atualização e o fato de a gramática eletrônica permitir a participação ativa do usuário, que pode reformular e organizar seu conteúdo, enviando seus comentários para a equipe de criação. Por fim, essa gramática pode também conter a descrição das regras, prover relevantes informações sobre o ensino de inglês e contar com a participação dos usuários em um processo de reestruturação constante.

### **3. Os prós e contras da linguagem do hipertexto**

A linguagem do hipertexto apresenta qualidades e deficiências que precisam ser ponderadas quando se pretende usá-la para a disponibilização de conteúdos na Internet. Vora e Helander (1997: 879-80) apontam, como ponto positivo, o potencial dessa linguagem na expressão de conteúdos e na sua capacidade de agilizar o acesso à informação, facilitar a recepção e a produção de textos e otimizar a apresentação e a manutenção dos conteúdos, ou bits de informação. Destacam as seguintes facilidades:

- acesso a informação feita através de associação;
- acesso a grande quantidade de conteúdos, compostos por várias mídias e conectados de formas várias;
- acesso ao conteúdo a partir de diferentes perspectivas e através de diferentes estratégias de busca;
- especificação de *links* e textos adicionais quando conveniente e relevante para satisfazer as necessidades de expressão;
- publicação de conteúdos, que se torna menos dispendiosa se comparada à publicação tradicional.

Mas, como já pontuado na seção anterior, alertam também para o fato de que o hipertexto apresenta desvantagens, sobretudo as relacionadas à navegação pelos *links*:

- o autor do hipertexto pode ter problemas para criar e manter as trilhas dos nós de informação e seus *links*;
- o leitor pode ter dificuldades em decidir qual *link* seguir diante de um grande número de opções, o que dele demanda esforço e concentração adicionais;
- o leitor pode sofrer de desorientação ou perda do senso de direção em um documento não-linear, afinal, o hipertexto possibilita a formação de redes de informação, mas não ensina como interpretá-la.

#### 4. A linguagem do hipertexto em ação

A grande popularidade e potencialidade de uso do hipertexto e da hipermídia na produção de conteúdos para a Web têm revelado que, mesmo com todo incentivo, e até mesmo fascínio que as potencialidades dessas linguagens exercem sobre os autores do século XXI, pouca atenção tem sido dada à proposição de princípios para nortear a elaboração do hipertexto e da hipermídia. Hemard (1997) fornece alguns subsídios que possibilitam o encaminhamento dessa questão. Propõe um conjunto de princípios de produção e fornece orientações para as etapas de pré-produção e de produção, com o objetivo de aliviar a complexidade das questões que surgem na autoria de hipertextos.

O autor afirma que o conjunto de princípios destaca conceitos importantes que formam a base do processo de produção de textos em uma infinidade de situações. Já o conjunto de orientações, proveniente de recente área do conhecimento que investiga a interação homem-máquina, foi selecionado através de processo de eliminação baseado em sua inaplicabilidade, duplicidade ou sobreposição.

A observância destes princípios é essencial para a produção de conteúdos para a Web:

- as tarefas projetadas devem ser compatíveis com os recursos tecnológicos disponíveis e com a experiência e habilidades da equipe envolvida na produção;
- o conteúdo disponibilizado deve ser disposto de modo adequado e claro (sintaxe); os comandos e ícones devem ser legíveis (semântica);
- a relação entre o comando e a imagem dos ícones deve ser transparente, isto é, as metáforas icônicas;
- o conteúdo expresso segundo a linguagem da Web não deve exigir esforço cognitivo adicional do leitor;
- o ambiente virtual que emoldura a informação deve apresentar consistência em termos do uso de metáforas e analogias (isto é, simulacros da realidade construídos no e pelo computador) relevantes para desmistificar o sistema e auxiliar o usuário na compreensão da complexidade da interface;
- o ambiente computacional deve prever Ajuda *on-line*.

Ao propor orientações, Hemard focaliza duas etapas: pré-projeto e projeto. Na primeira, enfatiza:

- a escolha de uma abordagem que potencialize as características e portabilidade do conteúdo a ser expresso no hipertexto;
- a análise das potencialidades e limitações das ferramentas computacionais disponíveis;
- a disponibilização de metáforas que os leitores possam reconhecer e que reflitam a natureza do hipertexto proposto;
- a possibilidade de domínio e controle do ambiente criado pelo leitor;
- a adoção da abordagem *top down*, isto é, do “global para o específico”;
- a elaboração conceptual de todo o hipertexto;
- a parcimônia na escolha de recursos gráficos e de animação, de fontes e de cores.

Na segunda, destaca:

- o estabelecimento de partições na disposição do conteúdo de modo a criar áreas claras e consistentes para a sua disponibilização;
- a criação adequada de telas para emoldurar os diferentes gêneros (epistolar, narrativo, forense, jornalístico etc.);
- a padronização de toda a metainformação que constitui o ambiente como, por exemplo, os menus, os comandos, os avisos, a ajuda etc.;
- o controle do número de *links* em cada página.

#### 4.1 Uma ilustração

Para ilustrar o procedimento de montagem de um hipertexto, esboça-se, na seqüência, na Fig. 1, o planejamento de um hipertexto, cuja finalidade é apresentar ao leitor os marcadores discursivos (doravante MDs) do inglês e do português (Schourup, 1999; Fraser, 1999; Bechara, 2000). Conforme especificado na figura, a partir da página de apresentação (página O), o leitor poderá saltar para qualquer uma das seguintes páginas: página DM, página E e página T. Se escolher a primeira opção, terá acesso aos MDs e, a partir dessa página, poderá saltar para a coleção de páginas dos marcadores que relacionam mensagens (páginas dmm), dos marcadores que relacionam tópicos (páginas dmt) e dos casos de ambigüidade (páginas dma). Cada uma dessas coleções de páginas oferece *links* para os MDs, apresentados nas páginas tm, tt e ta em forma de tabela. Se escolher a segunda opção, terá acesso à página de exercícios (E), que oferece *links* para a coleção de páginas em que os exercícios estão divididos de acordo com a tipologia dos MDs: exercícios de marcadores que relacionam mensagens (exe\_dmm), que relacionam tópicos (exe\_dmt) e os casos de ambigüidade (exe\_dma). Por fim, se escolher a terceira opção, terá acesso à página com os títulos dos textos (T) e, a partir desta, acesso à coleção de páginas em que estão os textos e aos *links* externos para a Web.

<p>Página 00 – Apresentação (O)</p> <p>Descrição: Página inicial de apresentação da estrutura e dos objetivos do hipertexto.</p> <p>Links: Os links apontam para as páginas Marcadores Discursivos (DM), Exercícios (E) e Textos (T).</p>
<p>Página 1 – Marcadores Discursivos (DMs)</p> <p>Descrição: Esta página contém o título Marcadores Discursivos e apresenta a classificação geral dos tipos de marcadores.</p> <p>Links: Os links apontam para a coleção de páginas de cada tipo de MD (&lt;dmt&gt;, &lt;dmm&gt;) e para os casos de ambigüidade (&lt;dma&gt;).</p>
<p>Coleção de Páginas &lt;dmt&gt;, &lt;dmm&gt; e &lt;dma&gt;</p> <p>Descrição: Estas páginas contém o título de acordo com o tipo e função de cada MD, com exemplos de uso. Há exemplos de MDs que podem ser fonte de ambigüidade.</p> <p>Links: O link externo aponta para a coleção de páginas (&lt;ex. 1-6&gt;).</p>
<p>Coleção de Páginas &lt;tt&gt;, &lt;tm&gt; e &lt;ta&gt;</p> <p>Descrição: Estas páginas contém tabelas de exemplos adicionais dos MDs de acordo com sua classificação. Há exemplos de MDs que podem ser fonte de ambigüidade.</p> <p>Links: Não há links.</p>
<p>Página 2 – Exercícios</p> <p>Descrição: Esta página leva o título Exercícios.</p> <p>Links: O link externo vai para a coleção de página &lt;ex.1-6&gt;.</p>
<p>Coleção de Página &lt;ex. 1-6&gt;</p> <p>Descrição: Estas páginas contém os exercícios dos MDs e estes, por sua vez, estão divididos de acordo com as funções dos MDs.</p> <p>Links: O link externo aponta para a coleção de páginas &lt;dmt&gt;, &lt;dmm&gt; e &lt;dma&gt;, e para a coleção de páginas &lt;txt 1-5&gt;.</p>
<p>Página 3 – Textos</p> <p>Descrição: Esta página contém o título Textos e a relação dos textos disponíveis no módulo.</p> <p>Links: O link aponta para a coleção de páginas &lt;tx. 1-5&gt;) em que se encontra cada um dos textos.</p>
<p>Coleção de Páginas &lt;tx. 1-5&gt;</p> <p>Descrição: Estas páginas contém os textos.</p> <p>Links: O link aponta para a coleção de páginas &lt;le. 1-5&gt;.</p>
<p>Coleção de Páginas &lt;le. 1-5&gt;</p> <p>Descrição: Esta coleção de páginas contém os links para a WEB, de acordo com cada tema tratado nos textos.</p> <p>Links: Os links externos apontam para a Web.</p>

Figura 1: Esquema de descrição da arquitetura do hipertexto..



A implementação do hipertexto pode ser feita de três modos diferentes: “traduzir” o esquema apresentado na Fig. 1 em hipertexto por meio da linguagem HTML (*HyperText Markup Language*), digitar toda a informação diretamente em uma interface que se encarrega de fazer a tradução automaticamente ou digitar a informação em um processador de textos convencional e, posteriormente, converter os arquivos em hipertexto por meio de um programa conversor.

Mesmo sendo a primeira alternativa a menos adequada, pois exige aprendizado da linguagem do hipertexto, experienciar o processo de escrever o hipertexto em HTML permite que o autor apreenda e avalie o grau de detalhamento e precisão necessários para que o hipertexto possa ser processado pelo computador. Essa tarefa minuciosa e analítica constitui também um momento ímpar para refletir o grau de coerência do texto objeto, posto que a explicitação de sua estrutura lógico-conceitual local e de sua macroestrutura textual global é tarefa essencial e pressuposta para que a conversão em HTML possa ser realizada com sucesso.

---

<html>... </html> sinaliza o início e o fim de codificação em html;  
<head> ... </head> sinaliza o início e o fim do cabeçalho;  
<title> ... </title> sinaliza o início e o fim do título;  
<body> ... </body> sinaliza o início e o fim do corpo de texto da página;  
<p> indica parágrafo;  
<ul> ... </ul> sinaliza o início e o fim de uma lista;  
<li> sinaliza itens de uma lista;  
<hr> sinaliza que o que vem a seguir deve ser tratado como um *link*.

---

Figura 2: Exemplo das funções básicas da linguagem HTML.

Como sinalizado, a conversão de um texto em língua natural para o formato em HTML exige domínio dessa linguagem, que se estrutura em torno de funções, parte das quais esta ilustrada na Fig. 2. A Fig. 3 ilustra a codificação da estrutura da página inicial da Capes<sup>6</sup> (cf. <http://www.capes.gov.br/>), com destaque para quatro funções descritas na Fig. 2.

<sup>6</sup> Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

```
<HTML>
  <HEAD>
    <link rel="SHORTCUT" href="/Imagens/Conteudo/LogoCAPES.ico">
      <TITLE>CAPES / MEC</TITLE>
    <META HTTP-EQUIV="Content-Type" CONTENT="text/html; charset=iso-8859-1">
  </HEAD>
  <FRAMESET ROWS="100,1*" FRAMEBORDER="NO" BORDER="0" FRAMESPACING="0">
    <FRAME NAME="Superior" SCROLLING="NO" noresize SRC="/Superior.htm"
  MARGINWIDTH="0" MARGINHEIGHT="0" FRAMEBORDER="NO" >
    <FRAMESET COLS="200,1*" FRAMEBORDER="NO" BORDER="0" FRAMESPACING="0">
      <FRAMESET ROWS="100%,1" FRAMEBORDER="NO" BORDER="0"
  FRAMESPACING="0">
        <FRAME NAME="Lateral" SCROLLING="NO" noresize
  SRC="/menPrincipal.htm" FRAMEBORDER="NO" MARGINWIDTH="0" MARGINHEIGHT="0">
          <FRAME NAME="Esc" SCROLLING="NO" NORESIZE SRC="" frameborder="NO"
  marginwidth="0" marginheight="0">
            </frameset>
          <FRAME NAME="Principal" SRC="/Destques.htm" FRAMEBORDER="NO"
  MARGINWIDTH="0" MARGINHEIGHT="0">
            </frameset>
          </frameset>
        </frameset>
      </frameset>
    </frameset>
  <NOFRAMES>
    <BODY BGCOLOR="#FFFFFF">
      </BODY>
    </NOFRAMES>
  </HTML>
```

Figura 3: Codificação em HTML da página inicial da Capes.

## Conclusões

Ao discutir questões relativas ao processamento automático de línguas naturais, e principalmente a necessidade de se construírem pontes entre os cientistas da linguagem e os cientistas da computação (para que, do trabalho colaborativo e cooperativo, seja possível a proposição de abordagens lingüístico-computacionais que melhor aproveitem os recursos expressivos desse novo meio de publicação, armazenamento e disseminação da informação em língua natural), Dias-da-Silva (1996; 1998) fatora a sistematização dos trabalhos de desenvolvimento de projetos dessa natureza em três "domínios" solidários e complementares: o lingüístico-expressivo, o computacional e o implementacional. Transposta essa abor-

dagem para este trabalho, foi possível equacionar, de maneira análoga, o processo de desenvolvimento de projetos de publicação de conteúdos na Web.

No primeiro domínio, procede-se à seleção dos conteúdos lingüístico e extra-lingüístico a serem veiculados pela Web. No segundo domínio, formalizam-se esses conteúdos em termos computacionais, visando à compatibilização entre as representações lingüisticamente motivadas e à necessária representação computacional, expressa, neste contexto, por meio da linguagem da Web, a HTML. No terceiro domínio, implementam-se essas representações em uma máquina específica, que passa a ser o *servidor* dos recursos e da informação.

Em particular, este trabalho pretendeu desvendar, sobretudo para o especialista em Letras, alguns meandros da Web, apresentando-lhe uma parte de sua história e de seus recursos, mostrando-lhe sua faceta gráfica e sua linguagem e, finalmente, ilustrando o uso de um de seus recursos essenciais: a montagem de um hipertexto.

**ABSTRACT:** *The Internet and its resources have been explored by several sectors of the global information society. This work aims at unveil, for the Humanities, part of what underlies this intricate virtual Web, fascinating for some, pernicious for others and spell-bound for the ones that investigate its intricacies. First, it is presented an overview of the Internet and its graphical facet: the Web. Then, it focuses on the potential of such language and information megastructure provides for research. Finally, it spots some conclusions and recommendations that should guide content publication in its hypertext markup language.*

**KEYWORDS:** *Internet; Web; hypertext; HTML; language.*

## BIBLIOGRAFIA

- CHENG, J.; MALAIKA, S. (Ed.). (1997) *Web gateway tools*. Lancaster: John Wiley; Sons, Inc.
- BARNES, S. (1994) Hypertext literacy. *IPCT – Interpersonal computing and technology: an electronic journal for the 21<sup>st</sup> century*, Washington, v. 2, n. 4, p. 24-36.
- BECHARA, E. (2000) *Moderna gramática portuguesa*. Rio de Janeiro: Lucena.
- DIAS-DA-SILVA, B. C. (1996) *A face tecnológica dos estudos da linguagem: o processamento natural das línguas naturais*. Araraquara, 272p. Tese (Doutorado). Faculdade de Ciências e Letras, Universidade Estadual Paulista. cap. 3.
- \_\_\_\_\_. (1998) Bridging the gap between linguistic theory and natural language processing. In: CARON, B. (Ed.). *International Congress of Linguistics*, 16, Oxford. *Proceedings [...]* Oxford: Pergamon-Elsevier Science. CD-ROM.
- DUNCAN, B. J. (1997) *Hypertext and education: (Post?)structural transformations*. [On-line]: <[http://www.ed.uiuc.edu/PES/97\\_docsduncan.html](http://www.ed.uiuc.edu/PES/97_docsduncan.html)> Acesso em 20.mar.2000
- FRASER, B. (1999) What are discourse markers? *Journal of Pragmatics*, v. 32, p. 913-52.
- GLOBAL REACH (Ed.). (2000) *Global Internet Statistics*. San Francisco. [On-line]: <<http://www.euromktg.com/globstats/index.html>> Acesso em 15.mar.2000
- HEMARD, D. P. (1997) Design principles and guidelines for authoring hypermedia language learning applications. *System*, v. 25, n. 1, p. 9-27.
- INTERNET SOCIETY (2000a) What is the Internet? Geneva. [On-line]: <<http://www.isoc.org/internet>> Acesso em 15.mar.2000
- \_\_\_\_\_. (2000b) A brief history of the Internet. Geneva. [On-line]: <<http://www.isoc.org/internet/history/brief.html>> Acesso em 15.mar. 2000
- LEVY, P. (1997) *As tecnologias da inteligência: o futuro do pensamento na era da informática*. Rio de Janeiro: Editora 34.
- LUFT, A. (2000) Estatística: internet tem 1 bil de páginas. *Folha de S. Paulo*. São Paulo, 26 de jan. Caderno 6, p. 4.
- MCINTYRE, D. R.; WOLFF, F. G. (1998) An experiment with WWW interactive learning in university education. *Computer; education*, v. 31, p. 255-64.

- NIB – NÚCLEO DE INFORMÁTICA BIOMÉDICA (2000) Aplicação de computadores no ensino médico. Recursos didáticos: ferramentas. Campinas: Unicamp. [On-line]: <[http://www.nib.unicamp.br/recursosferram\\_internet/internet.htm](http://www.nib.unicamp.br/recursosferram_internet/internet.htm)> Acesso Caderno 6
- PBS ONLINE (2000) Life on the Internet. Timeline. Alexandria. [On-line]: <<http://www.pbs.org/internet/timeline.html>> Acesso em 15.mar. 2000
- RNP – REDE NACIONAL DE PESQUISA (2000) O projeto Internet2. Rio de Janeiro. [On-line]: <<http://www.rnp.br/rnp2/rnp2-internet>>
- SCHOURUP, L. (1999) Discourse markers: tutorial overview. *Lingua*, v. 107, p. 227-65.
- SENGUPTA, S. (1996) Creating hypertext database to help Hong Kong English teachers teach writing. *System*, v. 24, n. 2, p. 187-98.
- VORA, P. R.; HELANDER, M. G. (1997) Hypertext and its implications for the internet. In: HELANDER, M.; LANDAUER, T. K.; PRABHU, P. (Eds.). *Handbook of human computer interaction*. Amsterdam: Elsevier Science B. V. p. 877-914.
- WARSCHAUER, M. (Ed.). (1996) *Virtual connections*. Honolulu: Univ. of Hawaii Press.
- WU, K.; TSUI, A. B. M. (1997) Teacher's grammar on the electronic highway: design criteria for telegram. *System*, v. 25, n. 2, p. 169-83.