

# O TEXTO EM SUPORTE ELETRÔNICO

(Electronic Texts)

Vilson J. LEFFA

*(Universidade Católica de Pelotas)*

O texto, para ser lido, precisa de um suporte onde ele possa ser de alguma maneira apoiado para a melhor visualização por parte do leitor. O suporte mais comum é o papel, mas muitos outros também podem ser usados como o plástico, o vidro, o acrílico, a madeira, o muro, a camiseta, o tronco de uma árvore e até suportes menos comuns e mais efêmeros como a areia da praia ou a fumaça de um avião no céu. Em todos esses casos, a visualização se dá de modo direto, sem o acréscimo de um processamento entre o suporte e os olhos do leitor geralmente basta a presença da luz.

Com as novas tecnologias de impressão, no entanto, há casos em que a simples presença da luz não é suficiente; é necessário um processamento adicional a fim de tornar o texto visível para o leitor. É o que acontece, por exemplo, com o texto informatizado: os bits, as unidades mínimas de informação compostas apenas de partículas de luz e escuridão, são totalmente invisíveis ao olho humano. É preciso um complexo sistema de processamento que não apenas amplie esses sinais, mas que também os traduzam em letras e gráficos para a compreensão do leitor.

O objetivo deste trabalho é descrever uma experiência de produção de texto informatizado. Para isso, mostra-se o que foi a criação do CD-ROM TELA (Textos em Lingüística Aplicada), destacando-se, entre outros aspectos, as diferenças que existem entre o texto impresso em papel e o texto em suporte eletrônico, as vantagens e desvantagem de cada mídia, os recursos da mídia eletrônica e as implicações que todos esses aspectos trazem para a produção e recepção do texto informatizado.

## APRESENTANDO TELA

TELA é um CD-ROM de Textos acadêmicos nas áreas de Linguística e Linguística Aplicada, com ênfase maior em Linguística Aplicada. Reúne, entre outros, textos completos de teses de doutorado, dissertações de mestrado, anais de congressos, relatórios de pesquisa, relatos de experiências, currícula de professores e pesquisadores, periódicos e alguns livros. O CD reúne cerca de 45.000 páginas de textos o que equivale, em termos comparativos, a 225 volumes de 200 páginas impressas em papel.

Esses textos, por estarem colocados em suporte eletrônico, podem ser acessados por dois caminhos básicos: (1) pela estrutura hierárquica, seguindo o princípio da árvore invertida, em que cada escolha ramifica-se em outras escolhas até chegar ao texto final; ou (2) pela busca indexada, onde o usuário digita uma ou mais palavras, e o programa lista os textos que contêm a palavra desejada, na ordem determinada pelo usuário.

A Figura 1 mostra a tela do menu inicial da estrutura hierárquica, onde cada figura ou quadro representa um link para outras telas, que se abrirão em outras opções, até que o leitor chegue ao texto final desejado.

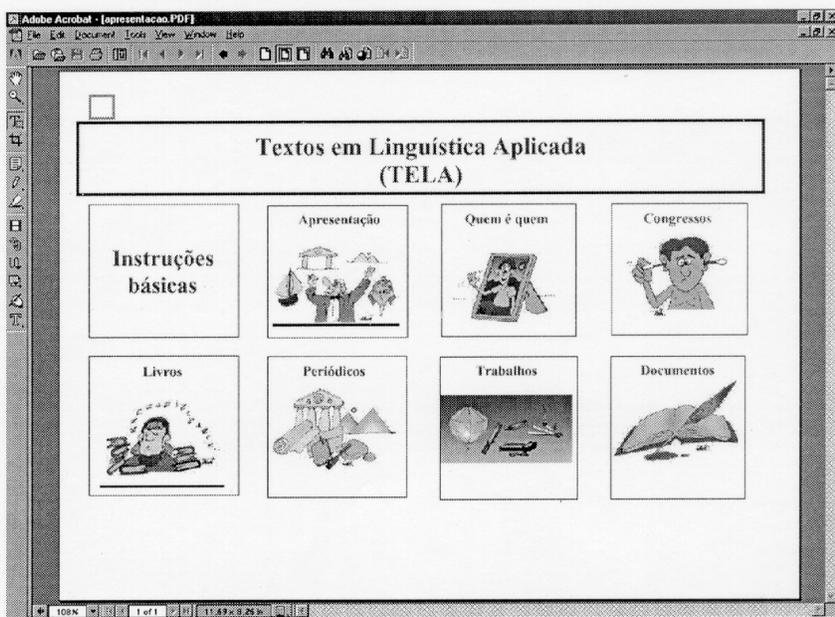


Figura 1 – Menu inicial de TELAR

Clicando em “Instruções Básicas”, por exemplo, o usuário chega a um outro menu, mostrado na Figura 2, onde se mostram quais são os primeiros passos que o usuário deve dar, como deve fazer para pesquisar na base de textos, como ver um resumo dos comandos, como contribuir, como citar e como tirar dúvidas através de perguntas e respostas. Em poucos segundos, qualquer uma das 45.000 páginas pode ser acessada com alguns cliques do mouse. Chegando ao texto final, há ainda uma lista de marcadores (bookmarks), que desce automaticamente à esquerda do texto para facilitar o acesso a qualquer uma de suas partes, permitindo uma navegação mais rápida e dando ao usuário não só a visão da página que está sendo lida mas também das partes principais do texto.

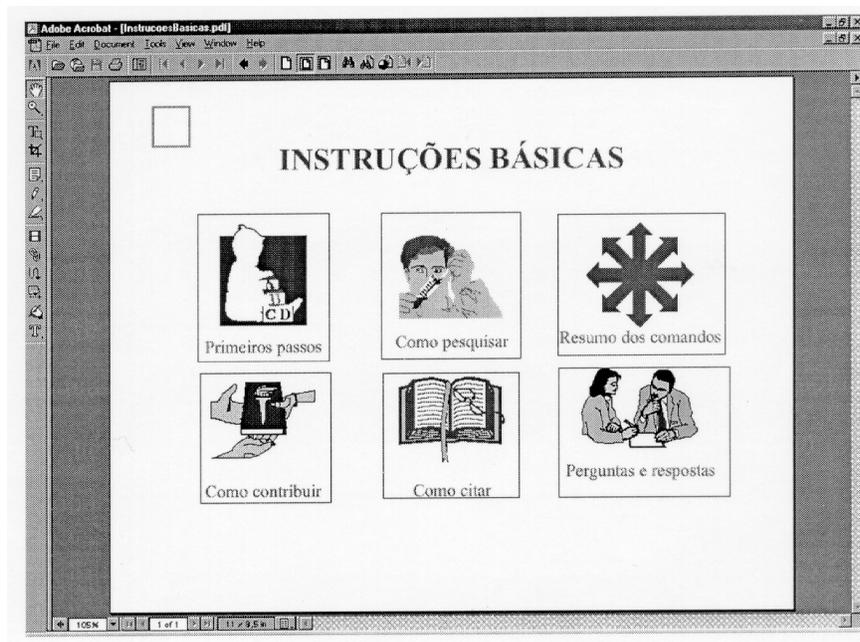


Figura 2 – Menu de instruções básicas

Além da facilidade de navegação dentro da estrutura hierárquica, há também a preocupação de tornar a interface com o programa o mais interativa possível, procurando-se até prever alguns movimentos incorretos do usuário. A Figura 3, por exemplo, mostra um desvio, que é automaticamente acionado, caso o usuário cometa o engano de clicar numa imagem exibida anteriormente, em vez de clicar na barra de ferramentas.

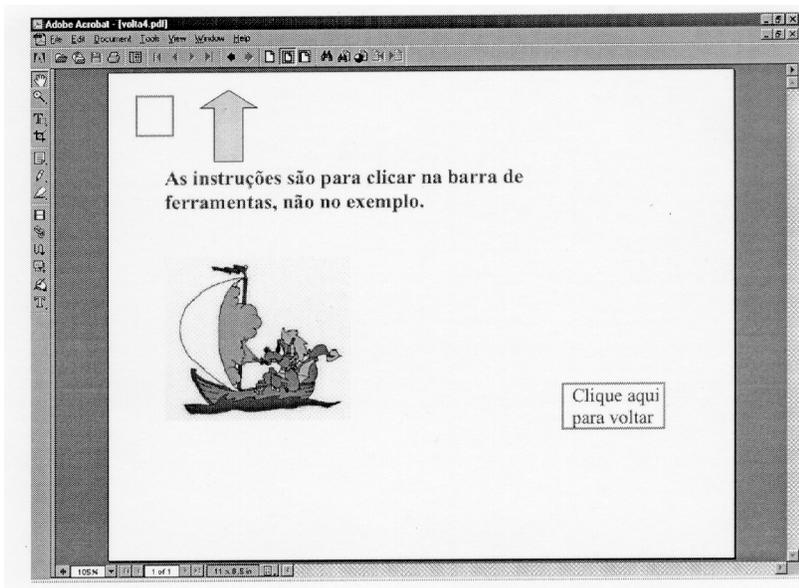


Figura 3 – Desvio acionado automaticamente quando o usuário comete um engano.

Os textos apresentados em TELA, no entanto, usam também os recursos da informática, permitindo que o usuário não apenas navegue pelo CD, como se ele fosse um jornal ou uma revista, mas que encontre rapidamente nas 45.000 páginas somente aquilo que lhe interessa. Para isso todos os textos estão indexados, possibilitando o acesso a qualquer trabalho publicado, seja por autor, assunto, tabela, gráfico, ou mesmo por qualquer palavra de qualquer texto. Basta clicar no ícone “busca”, na barra de ferramentas, e depois digitar a palavra ou palavras que deseja encontrar na caixa de diálogo que se abre. Imediatamente todos os textos que contêm as palavras digitadas começam a aparecer na tela. A Figura 4 mostra um exemplo em que o usuário pede para que sejam mostrados os textos que contenham as palavras “identidade” e “heterogeneidade” segundo o critério de “proximidade” o que significa que o programa mostrará em primeiro lugar os textos em que as duas palavras estejam mais próximas uma da outra. O programa permite também a busca de parte de palavras (ex.: todas as palavras que começam com “hiper-” ou terminam em “-logia”) e de segmentos maiores que a palavra (ex.: um sintagma como “análise de discurso”), permitindo, assim, ampliar ou afunilar qualquer busca.

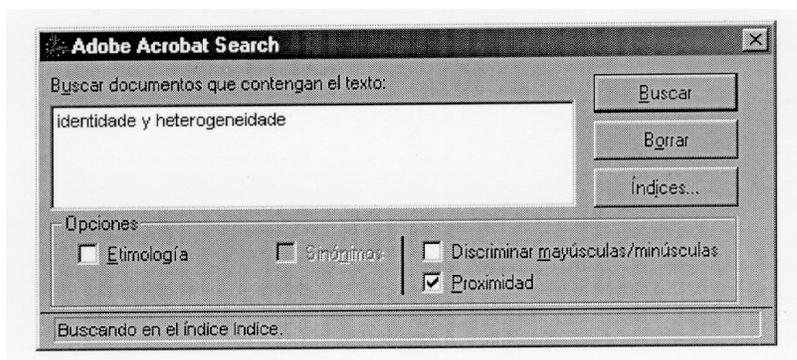


Figura 4 – Caixa de diálogo para busca indexada

TELA é, um grande projeto coletivo, feito com a colaboração de centenas de autores que disponibilizaram seus textos o que resultou num acervo de mais de 2.000 trabalhos individuais. Tem a peculiaridade especial de ser uma obra em que a maioria de seus leitores é provavelmente formada pelos autores que contribuíram com seus trabalhos.

## A OPÇÃO PELO SUPORTE ELETRÔNICO

O CD-ROM, como suporte textual, é uma mídia voraz, capaz de compactar numa superfície de 12cm de diâmetro, mais de 60.000 páginas típicas de texto acadêmico, incluindo fotos, gráficos e tabelas. O gigantismo inerente da mídia eletrônica assume proporções de portento quando se compara, por exemplo, o suporte eletrônico com o suporte em papel: o que cabe num CD ocupa o equivalente a 300 livros de 200 páginas. Comparado com o papel, o CD tem vantagens enormes, principalmente em termos de capacidade de armazenamento, mas há também algumas desvantagens.

Entre as desvantagens, as duas mais citadas são a falta de portabilidade e a baixa definição de imagens do texto eletrônico, quando exibido na tela do computador. O texto eletrônico realmente não dá para ser confortavelmente lido numa banheira ou numa fila de check-in para pegar um cartão de embarque num aeroporto mesmo com o uso de notebooks cada vez mais leves e potentes. Um livro, o leitor abre e começa a ler; um computador, o usuário liga e espera até que o processo de inicialização se complete,

o que às vezes pode levar mais de um minuto. A definição da página impressa em papel também é superior a da tela do computador. No mundo digital, ao contrário do analógico, as letras e imagens são literalmente desconstruídas em pontos para depois serem reconstruídas pelo usuário. Esses pontos, que deveriam ficar abaixo da percepção dos olhos, são às vezes claramente visíveis, e a definição é prejudicada.

Essas desvantagens também podem ser mais aparentes do que reais. Pode-se argumentar, por exemplo, em termos de portabilidade, que é muito mais fácil transportar um CD, mesmo dentro de uma caixa acrílica tradicional (na verdade pode ser acondicionado até dentro de um envelope plástico de espessura mínima) do que carregar 300 livros. Além do mais, o conteúdo de um CD, por ser uma informação digitalizada, pode ser transmitido de um ponto a outro do Planeta, em um tempo e a um custo mínimos, comparado ao que seria necessário para transportar 300 livros. Eu, pessoalmente, doaria com prazer todos os meus livros, periódicos e teses, que enchem várias prateleiras de minha pequena biblioteca, a quem fosse capaz de digitalizá-los para que eu pudesse colocá-los todos dentro do computador. Pouparia talvez os romances e poemas para, de vez em quando, ler um deles à sombra de uma árvore mas o resto desapareceria das prateleiras. Não só ganharia em espaço e portabilidade, como teria tudo automaticamente indexado, permitindo que eu logo achasse qualquer tópico ou referência, sem perda de tempo o que normalmente não me acontece quando procuro um texto na prateleira.

O problema da definição gráfica é também relativo; dependendo do tamanho e qualidade do monitor, muitas imagens podem ter uma aparência melhor na tela do computador do que no papel, às vezes até mais coloridas e brilhantes, independentes da luz ambiente. Por ser um texto líquido, o texto informatizado pode também ser facilmente ampliado ou reduzido, sem perda da definição, já que normalmente usa técnicas de vetorização.

As vantagens do texto impresso em papel parecem depender mais de fatores circunstâncias (ex.: ler na cama para esperar o sono) do que das qualidades intrínsecas do suporte textual. Quando se lê um texto acadêmico, cujo objetivo é essencialmente prático, envolvendo a pesquisa e uma participação mais ativa do leitor, que muitas vezes lê um texto para aproveitá-lo no que está pesquisando, o suporte eletrônico oferece enormes vantagens sobre o texto em papel.

A maior delas é que o suporte eletrônico dilui a fronteira entre a leitura e a escrita. Como o espaço usado para ler e escrever é o mesmo a tela do computador e como o texto é líquido, a diluição do que se lê no que se escreve dá-se de modo natural, limitada apenas pela capacidade do autor em aproximar textos compatíveis. A facilidade da reescrita, montagem e composição elimina o plágio e constrói a autoria, removendo as costuras entre o já-dito e o não-dito.

## A ESCOLHA DO SOFTWARE

Uma decisão importante na criação do Projeto TELA, foi a escolha do software a ser usado para gerenciar o acervo de textos. Três critérios básicos orientaram essa escolha: idealmente o software deveria ser (1) genérico, isto é, disponível na maioria dos computadores e sem custos adicionais para o usuário final; (2) não exigir instalação no computador para exibição dos textos, podendo ser lido diretamente do CD; e (3) mostrar os textos no formato original.

Inicialmente pensou-se em fazer a conversão dos arquivos coletados para arquivos do tipo “.html”, que podem ser lidos por qualquer “browser” (Netscape, Internet Explorer, etc.). Os principais “browsers” estão já instalados na maioria dos computadores, são distribuídos gratuitamente e podem ler os textos diretamente do CD. Têm, no entanto, o inconveniente de alterar a formatação original dos textos não só durante a conversão, mas também de um computador para outro, fazendo adaptações para cada configuração diferente de vídeo.

Depois de várias tentativas e ensaios, optou-se pelo Acrobat Reader da Adobe Systems. Não é de uso tão generalizado como os “browsers”, mas atende aos três critérios exigidos. Pode ser incorporado aos “browsers” como “plug-in”, lê os textos diretamente do CD, é amplamente usado em manuais e divulgação de produtos e, principalmente, mantém a formatação original dos textos, mostrando-os tais como foram impressos, incluindo, figuras, numeração de páginas, notas de rodapé, etc. Incorpora ainda alguns recursos adicionais que se mostraram extremamente úteis, como um completo sistema de indexação e facilidade de navegação pelo acervo de textos.

Acrobat é apresentado em duas versões: uma completa (Adobe Acrobat) para a conversão dos textos em arquivos “PDF” (Portable Document Format) e uma versão somente para leitura (Acrobat Reader). O Adobe Acrobat produz documentos que podem ser lidos por diferentes plataformas (Windows, Mac OS e Unix) sem mudar a aparência original do texto. O Acrobat Reader, por sua vez, apenas lê os textos produzidos pelo Adobe Acrobat, mas, ao contrário da versão completa, é distribuído gratuitamente pela Adobe Systems, podendo ser baixado pela Internet, o que facilita a publicação e distribuição dos textos. Acrobat é apresentado em várias línguas, inclusive em português, embora, neste caso, sem o mecanismo de busca. Uma vez convertido o arquivo, ele pode ser lido em qualquer versão do programa e em qualquer língua. Para melhor aproveitamento dos recursos usados na criação do Projeto TELA, no entanto, recomenda-se o uso de uma versão com o mecanismo de busca (espanhol, inglês, francês, etc.), inexistente em português.

## **A APRESENTAÇÃO DO ACERVO: DO ESTÁTICO AO DINÂMICO**

Uma das grandes preocupações na elaboração do projeto TELA foi como organizar e apresentar os textos na tela do computador facilitando ao máximo a leitura e a pesquisa nos milhares de páginas que compõem o acervo. O leitor deveria não só encontrar o que desejasse da maneira mais rápida possível, mas também não se desorientar dentro das inúmeras camadas de texto que formam hipertexto do Projeto TELA.

A metodologia básica consistiu, inicialmente, da coleta dos textos e, posteriormente, da organização desses textos.

A coleta dos textos partiu de uma chamada de trabalhos dentro dos moldes tradicionais, seguida de contatos pessoais com organizadores de congressos, editores de periódicos, autores de trabalhos, incluindo preferencialmente textos que já tivessem passado por algum processo de avaliação como teses de doutorado e dissertações de mestrado já defendidas. Na medida em que o acervo foi aumentando, possibilitando a demonstração do programa em congressos e encontros de professores, aumentou também a receptividade ao projeto. O salto mais rápido foi das 10.000 para as 40.000 páginas, com a incorporação, entre outros textos, dos anais do GELNE (Grupo de Estudos Lingüísticos do Nordeste), CELSUL (Círculo

de Estudos Lingüísticos do Sul), ABRALIN (Associação Brasileira de Lingüística), ASSEL (Associação de Estudos da Linguagem do Rio de Janeiro) e várias teses de doutorado.

Para a organização dessas dezenas de milhares de páginas experimentaram-se diferentes abordagens. Chegou-se à conclusão de que há duas estruturas hierárquicas possíveis para a apresentação de textos em suporte eletrônico: a hierarquia estática e a hierarquia dinâmica.

A hierarquia estática é aquela que caracteriza a apresentação de textos em papel, podendo obviamente também ser apresentada em suporte eletrônico. Segue o modelo de uma árvore invertida, partindo do tronco para as folhas, de acordo com a taxonomia pré-estabelecida pelo compilador dos textos. Quando o leitor chega para iniciar a consulta, a estrutura dos tópicos e as relações entre eles já estão estabelecidas, cabendo portanto ao leitor adaptar-se e orientar-se por essa estrutura para chegar ao que deseja.

A hierarquia dinâmica não é pré-estabelecida pelo compilador do acervo; estrutura-se e reestrutura-se a cada consulta que o leitor faz e só é possível em suporte eletrônico. Não há menus estáticos que inicialmente descem na tela para a escolha do leitor. A interação inicia-se por caixas de diálogo (“dialogue boxes”) e a relação menu/escolha é invertida: primeiro o leitor escolhe, digitando na caixa de diálogo o que deseja pesquisar; depois aparece o menu, estruturado a partir do que o leitor deseja. A grande vantagem dessa hierarquia dinâmica é que os textos mais pertinentes aos interesses do leitor são sempre colocados no topo.

O *Adobe Acrobat* apresenta uma abordagem para a hierarquização dos textos que é extremamente simples na sua concepção e ao mesmo tempo extremamente eficaz nos resultados, baseada no critério da proximidade: os textos em que as palavras escolhidas pelo leitor estão mais próximas uma da outra são mostrados em primeiro lugar. Além de sintagmas simples (“ensino”), é possível também usar sintagmas complexos (“ensino de línguas estrangeiras”) e mesmo parte de palavras (“ensin\*”, “\*eira”, “\*tran\*”) ou seja, segmentos iniciais, médios e finais de palavra. Isso permite não só refinar a pesquisa através de inúmeras combinações e chegar exatamente ao tópico que se deseja da maneira mais rápida possível, mas também possibilita a investigação de aspectos puramente lingüísticos, como listar as palavras do acervo que contenham um determinado prefixo, sufixo, radical, rima, etc.

Inúmeros outros projetos de acervos eletrônicos foram examinados, incluindo as versões digitais de periódicos acadêmicos (*Tesol Quarterly*; *Theory into Practice*), acervos de obras literárias (*Library of the Future*; *World Literary Heritage*), coleções de jornais e revistas (*CD-ROM Folha 97*; *Time Almanac of the 20th Century*; *Super Interessante*, *10 Anos de Revista em um CD-ROM*; *Exame*, *o Melhor dos Anos 90*), enciclopédias (*Compton's*, *Enciclopédia Interativa*, *Encarta*, *Infopedia*, *Barsa*), CD-ROM's de congressos (*Reuniões Anuais da SBPC*, *CELSUL*, *CELLIP*, *ABRALIN*, *ANPOLL*), bases de dados (*Eric*, *Sociological Abstracts*, *MLA*), etc. A maioria desses acervos apresenta uma combinação das duas estruturas hierárquicas (estática e dinâmica), mas muitos ainda usam programas específicos que precisam ser instalados no computador para que os textos possam ser vistos. Na medida em que se instala um CD depois do outro, cada um exigindo um programa diferente para ser lido, não só ocupando espaço no computador mas apresentando uma interface diferente e conseqüentemente exigindo uma aprendizagem do usuário, fica clara a vantagem de um programa genérico, comum para a leitura de todos os CDs. Essa foi a primeira conclusão a que se chegou depois da análise de todos esses projetos.

A segunda conclusão foi de que se deveria usar as duas hierarquias, tanto a dinâmica como a estática. Permite-se, assim, ao leitor percorrer os menus estáticos, "folheando" o acervo como se fosse uma revista ou livro, e dá-se também a oportunidade de usar as caixas de diálogo, possibilitando os menus dinâmicos por onde se chega exatamente ao que se deseja ler. Enfatiza-se, no entanto, nas instruções iniciais do Projeto que, para melhores resultados, o leitor deve usar a segunda opção.

O uso das duas hierarquias estática e dinâmica, combinado com detalhes acrescentados durante a compilação do acervo, dá ao Projeto TELA, várias opções de navegação, tentando manter o melhor equilíbrio possível entre o específico e o geral. Mesmo quando o leitor encontra especificamente aquilo que deseja, ele pode, querendo, ter uma visão mais geral de outros textos que estão ao redor daquele texto mostrado na tela. Ao ler uma tese de doutorado, por exemplo, o leitor tem na sua frente, simultaneamente, o texto que está lendo, o título do capítulo em destaque e a estrutura da tese com os títulos dos outros capítulos. Se o texto for um trabalho apresentado em congresso, ele tem uma visão dos outros títulos e pode chegar, com dois cliques de mouse, aos detalhes necessários para a referência bibliográfica do trabalho, apresentado como um decalque amarelo colado sobre o texto, sem afetar sua apresentação gráfica.

A decisão sobre o que expor na tela e o que ocultar quando o texto está sendo lido só foi tomada depois de várias tentativas, onde se tentou chegar ao máximo possível de informação sem congestionar a tela com muitos detalhes. Em princípio, observou-se que quanto mais informação, menos necessidade de aprendizagem por parte do leitor no uso do programa, mas mais congestionamento na tela. Para conseguir o equilíbrio entre essas três variáveis, optou-se às vezes por ocultar a informação acessória, deixando no entanto uma pista para que o texto oculto pudesse ser acessado.

A Figura 5 mostra parte de uma página com menos detalhes, no caso sem a sugestão de referência bibliográfica e sem o painel de navegação à esquerda. O leitor, querendo, clica sobre o decalque (de cor amarela na tela do computador) e a sugestão de referência bibliográfica aparece (Figura 6), permitindo, por exemplo, que o leitor cole o texto no trabalho que estiver digitando demonstrando também o que se falou acima sobre a diluição da fronteira entre o ler e o escrever. Se o leitor, por outro lado, clicar no ícone do painel de navegação, na barra de ferramentas superior, a coluna de marcadores desce à esquerda da tela, permitindo que o leitor visualize também as seções gerais dos anais do congresso onde está publicado o trabalho. Clicando no sinal positivo (+), os trabalhos de cada seção são mostrados, com destaque em inverso para o trabalho que está sendo lido, permitindo, assim, ao leitor uma visão geral dos trabalhos do congresso.

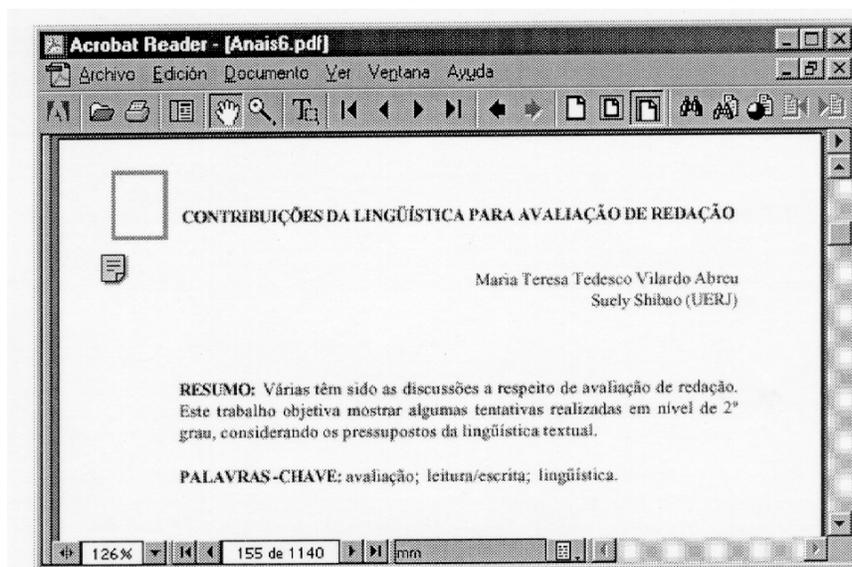


Figura 5 – Tela com menos detalhes

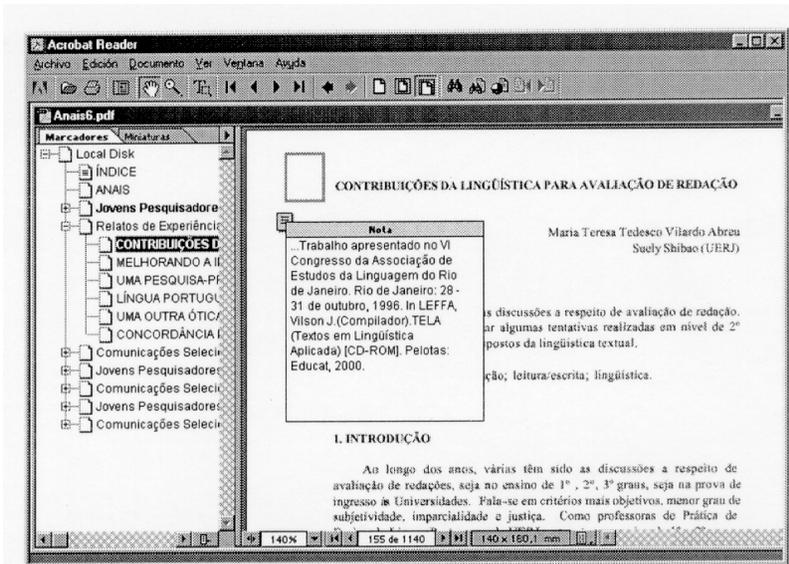


Figura 6 – Tela com mais detalhes

Outra decisão que só foi tomada depois de várias tentativas incluía a questão do que deveria ser inicialmente mostrado ao leitor. Havia dois extremos: (1) mostrar o máximo de detalhes, congestionando a tela, como na Figura 6, e deixando a limpeza por conta do leitor; ou (2) mostrando o mínimo de detalhes (Figura 5), para que o leitor expandisse como desejasse. Mais uma vez, optou-se pelo meio termo, retirando o que parecia obstruir a leitura, como texto sobre texto no caso da referência bibliográfica, e deixando a informação que não obstruía, como o painel de navegação à esquerda. Os textos, portanto, iniciam automaticamente neste meio termo, permitindo ao leitor aumentar ou diminuir o detalhamento conforme suas necessidades ou preferências.

## Os usos do Projeto Tela

TELA foi inicialmente projetado unicamente para pesquisa, facilitando o acesso a documentos que ficam muitas vezes restritos a bibliotecas locais, sem a divulgação que merecem. À medida, porém, que se começou a testar e a usar o projeto, descobriu-se que ele também pode ter outras finalidades, inclusive para levantamento de dados estatísticos. Entre essas

finalidades, destacam-se freqüência de citações, tendências da área e estudos lingüísticos.

É óbvio que a principal finalidade do Projeto TELA é a pesquisa bibliográfica; nenhuma outra fonte eletrônica no Brasil reúne num só lugar tanta informação sobre uma determinada área de conhecimento como o Projeto TELA. São dezenas de milhares de páginas totalmente indexadas e cobrindo praticamente tudo que interessa aos pesquisadores da área sob diferentes perspectivas, incluindo tópico (ex.: língua materna, língua estrangeira, análise de discurso, sintaxe, fonologia, estratégias de aprendizagem, etc.) e tipos diferentes de publicação (teses, livros, periódicos, anais, etc.). Apresenta não só as fontes onde os textos podem ser encontrados, mas os próprios textos, em versão integral, facilitando a citação, que nem precisa ser digitada basta copiar e colar. Além do texto para citação, a própria referência de muitos trabalhos é também facilitada, como mostra a Figura 6.

Na medida em que se testava o Projeto TELA, descobriu-se que ele também tinha inúmeras outras finalidades. Sendo um arquivo eletrônico, permite, com relativa facilidade, vários tipos de levantamentos de dados sobre os mais diversos tópicos. O Quadro 1 mostra alguns exemplos de possíveis levantamentos, incluindo tendências, autores, associações científicas e instituições de ensino superior. O escore apresentado neste quadro não indica um número absoluto (o total de citações no acervo), mas um número relativo ao número de segmentos existentes no projeto. Se um autor, por exemplo, for citado mais de uma vez em um mesmo trabalho, conta como uma única citação. O número máximo possível para cada item é 100. O quadro não revela necessariamente os itens mais citados em todo o acervo mas apenas os que foram pesquisados, partindo de itens previamente selecionados.

A preferência pelo uso de escores relativos é a maior validade dos resultados proporcionada por essa metodologia. Embora seja também possível fazer a pesquisa em termos absolutos, levantando o número total de citações em todo o acervo, o uso de escores relativos proporciona um resultado mais confiável. Isso pode ser percebido, por exemplo, verificando o número de ocorrências da palavra ABRALIN. Em termos absolutos ela é citada mais de 1.000 vezes, mas esse escore elevado deve-se ao fato de esta palavra fazer parte do rodapé do *Boletim da ABRALIN*, onde é citada em

**Quadro 1 – Exemplos de algumas citações levantadas  
do acervo do Projeto TELA**

<b>Tópicos</b>	<b>Escore relativo</b>
Aprendizagem de línguas	59
Interação em sala de aula	49
Sintaxe	40
Alfabetização	36
Aprendizagem do inglês	31
Aquisição da L2	30
Aprendizagem do espanhol	25
Reescritura	14
Aquisição da L1	10
<b>Nomes citados</b>	<b>Escore relativo</b>
Marilda Cavalcanti	30
Luiz Paulo Moita Lopes	27
Maria Antonieta Alba Celani	24
Ângela Kleiman	20
Francisco Gomes de Mattos	17
<b>Associações científicas</b>	<b>Escore relativo</b>
ABRALIN	19
ALAB	9
SBPC	9
ABRALIC	5
<b>Universidades</b>	<b>Escore relativo</b>
PUC-SP	58
UNICAMP	50
UFRJ	38
PUC-RS	35
UFMG	31
USP	30
UFRGS	27

cada uma das 867 páginas. Já em termos relativos, todas as citações que ocorrem no *Boletim da ABRALIN*, ficam reduzidas a apenas uma citação. O critério pode parecer demasiadamente rigoroso, mas na medida em que é usado com todos os itens, produz um resultado mais confiável, mostrando, por exemplo, que, em comparação com as outras sociedades científicas, a ABRALIN é mais citada no Projeto TELA.

Esses resultados são proporcionados automaticamente pelo programa, bastando que o usuário digite uma palavra ou combinação de palavras. É possível também uma contagem manual dos itens, avaliando quando a ocorrência deve ser ou não considerada uma citação (descartando, por exemplo, as ocorrências que apareçam em cabeçalhos ou rodapés). Essa metodologia, ainda que manual e mais demorada, é também facilitada pelo programa, na medida em que mostra automaticamente cada citação para a avaliação do pesquisador o que só é possível em texto com suporte eletrônico.

Finalmente, é também possível fazer uma pesquisa puramente lingüística do acervo que compõe o Projeto TELA, considerando aspectos lexicais, sintáticos, morfológicos, etc. ainda que o Projeto não tenha sido planejado com essa finalidade. Em um estudo da anáfora, por exemplo, pude verificar rapidamente a hipótese de que o sintagma nominal, quando ligado aos possessivos *seu, sua, seus, suas* pela conjunção coordenada *e*, é o antecedente desses possessivos (*a língua materna e seu ensino, a norma e sua violação, o amor e seus mistérios, a sociedade e suas tensões*). Posso verificar também que a expressão “análise *do* discurso” é 2,5 vezes mais freqüente do que “análise *de* discurso”.

O *Acrobat Reader* oferece inúmeros recursos de pesquisa, que possibilita inúmeras formas de busca. É possível pesquisar não só através da lógica booleana, usando diversas combinações de “e” e “ou”, mas também usando combinações de diferentes palavras e mesmo partes de palavras.

## CONCLUSÃO

O texto em suporte eletrônico oferece vantagens e desvantagens sobre o texto impresso em papel. Dependendo dos objetivos com que o texto é lido, essas vantagens podem ser maiores ou menores. No caso do texto acadêmico, cujo objetivo de leitura é principalmente a pesquisa biblio-

gráfica, as vantagens superam em muito as desvantagens, tanto em quantidade como em qualidade. Em termos de quantidade, o suporte eletrônico permite estocar, em um único CD, o equivalente a 60.000 páginas de texto; em termos de qualidade, permite que o texto seja totalmente indexado, de modo que qualquer tópico, autor ou mesmo palavra possa ser rapidamente acessado, facilitando a consulta. O uso de hierarquias dinâmicas, construídas durante a consulta, ao invés das hierarquias estáticas, sedimentadas durante a produção, pode dar ao leitor uma preciosa economia de tempo, na medida em que automaticamente reordena os textos por ordem de relevância, colocando em primeiro o que é mais pertinente à consulta do leitor.

Demonstraram-se, neste trabalho, algumas diferenças básicas entre o texto eletrônico e o texto impresso em papel. Para isso, usou-se como exemplo ao mesmo tempo em que se o descreveu o Projeto TELA (Textos em Lingüística Aplicada), que reúne em CD-ROM cerca de 45.000 páginas de texto nas áreas de Lingüística e Lingüística Aplicada. Mostraram-se as várias maneiras como o texto que compõe o acervo do Projeto na verdade um hipertexto pode ser consultado, com várias opções de navegação. Descreveu-se também o que foi a feitura do Projeto TELA, desde a coleta dos dados, reunindo mais de 2.000 trabalhos, até a organização e compilação final, incluindo algumas opções críticas como facilidade de uso, sem necessidade de aprendizagem pelo usuário, escolha de software genérico e apresentação dos textos em seu formato original.

Na medida que se foi compilando e testando o Projeto, descobriu-se que ele tem mais usos do que foi originalmente planejado, incluindo não apenas a pesquisa bibliográfica, mas também levantamentos estatísticos, já que inúmeras contagens podem ser automaticamente realizadas. Entre esses levantamentos, por exemplo, podem ser verificados quais são os autores mais citados, as tendências da área, as instituições de ensino superior e sociedades científicas mais freqüentemente mencionadas, etc. Mesmo pesquisas puramente lingüísticas podem também ser realizadas com os recursos de busca inseridos no programa que gerencia o acervo de textos do Projeto TELA.

Finalmente, um aspecto que surpreendeu e que também não estava planejado, foi a receptividade ao Projeto, durante sua compilação. TELA é uma obra essencialmente coletiva, feita da colaboração de centenas de autores que generosamente enviaram seus trabalhos.

Recebido em agosto de 2000