

(11) Número de Publicação: **PT 917698 E**

(51) Classificação Internacional:
G09B 17/00 (2006.01) **G06F 17/27** (2006.01)

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

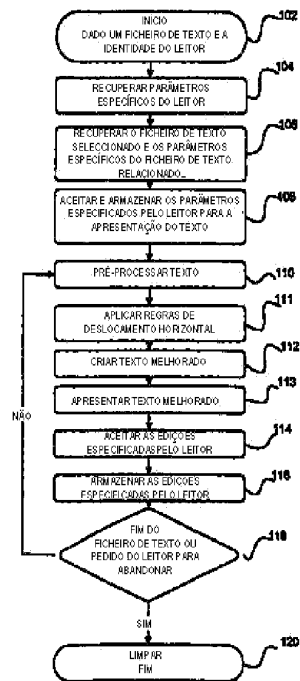
(22) Data de pedido: 1997.07.24	(73) Titular(es): RANDALL C. WALKER 923 TENTH STREET S.W. ROCHESTER, MN 55902 US
(30) Prioridade(s): 1996.08.07 US 693444	
(43) Data de publicação do pedido: 1999.05.26	
(45) Data e BPI da concessão: 2007.07.04 050/2007	(72) Inventor(es): RANDALL C. WALKER US
	(74) Mandatário: PEDRO DA SILVA ALVES MOREIRA RUA DO PATROCÍNIO, N.º 94 1399-019 LISBOA PT

(54) Epígrafe: **PROCESSADOR DE TEXTO.**

(57) Resumo:

RESUMO

"PROCESSADOR DE TEXTO"



Método e aparelho para o melhoramento de texto para a apresentação de texto para leitura humana melhorada. O método inclui extrair atributos específicos de texto a partir de texto legível numa máquina e variar a apresentação de texto de acordo com os atributos. A forma de realização preferida inclui extrair partes de fala e pontuação de uma frase, aplicar regras de dobragem que utilizam as partes de fala para determinar os pontos de dobragem, utilizar os pontos de dobragem para dividir a frase em segmentos de texto, aplicar regras (111) de deslocamento horizontal para determinar o deslocamento horizontal para os segmentos de texto e apresentar cada um dos segmentos (113) de texto, numa nova linha tendo o deslocamento

DESCRIÇÃO

"PROCESSADOR DE TEXTO"

A presente invenção refere-se a um método de processamento de texto para obter, a partir de um texto legível numa máquina, uma apresentação de texto melhorada baseada em parâmetros específicos do leitor, e a um processador de texto para executar este método.

A intenção da leitura é provocar um impacto na mente do leitor. Isto é verdade quer o texto a ser lido compreenda um romance ou um monitor de alerta de uma aeronave. O material apresentado num meio não-textual transporta informação bem adaptada para absorção humana para além do texto unidimensional correspondente. As imagens fixas apresentam atributos de objectos, tal como cores, tamanhos relativos, posições relativas, padrões, agrupamentos e hierarquias. Um objecto é visionado num contexto. Por exemplo, uma imagem de um pequeno objecto localizado juntamente com outros num canto transporta uma informação diferente da que o mesmo objecto, aumentado, sozinho, e no centro. As frases e parágrafos iniciais adicionam informação contextual unicamente pela sua posição. As imagens em movimento possuem mudanças de movimento e temporais como atributos adicionais. Muita informação chega ao cérebro visualmente e através de reconhecimento de padrões. Outra informação chega como informação áudio, transportando informação através do tom, da mudança de modulação da voz e da mudança do nível de volume.

A apresentação de texto em computador para leitura humana não tem conseguido aproveitar muito do cérebro humano. Como resultado, apenas uma parte da largura de banda possível é utilizada, e a apresentação de texto em computador é, muitas vezes, considerada como não amigável, constante, mecânica, e em geral, com algum tipo de falha. Quando lhes é dado a escolher, a maioria das pessoas prefere ler um livro impresso em vez de um livro num écran de computador. Actualmente, a apresentação de texto em computador é pobre relativamente a um livro, e ainda mais pobre relativamente ao seu potencial.

Foi feito algum trabalho na apresentação de texto computadorizado. O documento US-A-4880385 divulga um dispositivo de leitura rápida optomecânico que permite a visualização de texto impresso em papel através de uma janela de visualização, uma linha de cada vez, avançando as linhas automaticamente.

Foram feitos avanços em métodos para a análise computadorizada de linguagem natural dividindo-a em partes de fala. O documento US-A-5475588 divulga um sistema de análise melhorado para a criação de árvores de análise. O documento US-A-5331556 divulga um método para armazenar partes de informação de fala num ficheiro juntamente com o texto original para uma busca subsequente melhorada do texto. Os documentos US-A-4661924 e US-A-5146405 divulgam métodos para eliminar a ambiguidade de múltiplas partes de fala. Os documentos US-A-4887212, US-A-5068789 e US-A-5083268 divulgam métodos para analisar texto de linguagem natural. A leitura é um processo complexo e existem muitos factores que determinam diferenças no desempenho da leitura entre leitores e até para o mesmo leitor em ocasiões diferentes. Estes incluem condições neurofisiológicas inatas, tais como a dislexia, mas também a idade; factores

comportamentais e motivacionais; nível de educação e prática anterior de leitura; e limitações perceptuais. Um leitor também pode ter diferentes objectivos de leitura que afectarão a forma como ele/ela abordam o material de leitura. Apesar de tudo o indicado acima, o texto é apresentado por computadores, no pior caso, unidimensionalmente, como "pérolas de um colar", no melhor caso, de forma semelhante ao texto bidimensional apresentado pelos livros. De facto, se o texto é apresentado na forma impressa ou num monitor electrónico interessa pouco porque a apresentação do texto é essencialmente idêntica.

O documento US-A-5060155 divulga um método para codificar, de forma não ambígua, múltiplas análises de uma sequência de palavras de linguagem natural em gramática dependencial, na qual as dependências são definidas entre pares de palavras, cada par consistindo numa palavra superordenada ou governante e numa palavra relacionada ou dependente. Determina-se uma rede sintáctica que é representada como uma árvore consistindo de nós ligados mutuamente por arestas e compreendendo, pelo menos, um nó superior, um ou mais nós terminais e, eventualmente, vários nós intermédios, cada nó servindo de apontador se existir apenas uma alternativa e servindo de ponto de escolha se existirem várias alternativas, pelo que cada um dos nós apontadores é atribuído a uma palavra da sequência e cada aresta é atribuída à relação sintáctica entre os dois nós ligados por essa aresta, pelo que cada nó é codificado por um identificador que, no caso de um nó apontador, está directamente relacionado com a entrada de uma palavra no dicionário e, no caso de um ponto de escolha, compreende uma lista de identificadores adicionais, um dos quais tem de ser escolhido.

O documento US-A-4613309 divulga um modelo educacional para mostrar as partes da gramática de uma linguagem, de fala, sintaxe, e estrutura de frases nas quais trinta elementos individuais de fala são identificados e agrupados em dez conjuntos distintos que são materializados em modelos físicos. Cada modelo representa uma palavra ou grupo de palavras, referidas como "elementos", que formam os blocos lógicos de construção da linguagem. Certos elementos são, de facto, elementos híbridos uma vez que reflectem as propriedades e derivações únicas de verbos, frases e proposições. Utilizam-se símbolos para designar a função de cada elemento; a cor é utilizada para designar o conjunto do elemento, e a forma é utilizada para designar a parte da frase e parte de fala. O utilizador interliga cada elemento com outros elementos para formar uma frase, proposição ou oração como uma alternativa a um diagrama de frases convencional. Os modelos podem ser concretizados em forma impressa, plástico, vídeo, ou *software*.

São muitos os exemplos da constante apresentação de texto. As frases tópico são apresentadas da mesma forma que as outras frases num parágrafo. As palavras técnicas são apresentadas da mesma forma que as palavras não-técnicas. Os adjectivos aparecem da mesma forma que os substantivos. Os termos educacionalmente difíceis são apresentados da mesma forma que os termos simples. O último parágrafo num capítulo é apresentado da mesma forma que o primeiro parágrafo. O texto é apresentado com justificação à esquerda, com linhas partidas de largura constante, requerendo um movimento da vista reflexo da extrema direita para a extrema esquerda a intervalos regulares. O texto é partido a meio da oração, meio da frase e meio do pensamento, aplicando regras de composição antigas. Este tipo de texto força os olhos a deslocarem-se para trás e para a frente ao longo de grandes

distâncias, mimetizando o movimento do carro duma máquina de escrever. O texto é avançado manualmente, dividido em pedaços determinados por quantas linhas cabem no écran.

O documento US-A-5060155 mostra um método para codificar, de forma não ambígua, múltiplas análises de uma sequência de palavras de linguagem natural em gramática dependencial, na qual as dependências são definidas entre pares de palavras, cada par consistindo numa palavra superordenada ou governante e numa palavra relacionada ou dependente. Para cada palavra da sequência é determinado um índice de palavra, representando a posição da referida palavra na sequência. Todos os possíveis dependentes de cada palavra são determinados, bem como a relação entre a palavra e os dependentes utilizando um algoritmo de análise em combinação com uma gramática e um dicionário nos quais todas as palavras da linguagem estão armazenadas juntamente com a sua interpretação sintáctica e um índice de interpretação, representando a posição da interpretação sintáctica da palavra no dicionário de forma a distinguir entre múltiplas interpretações sintácticas da referida palavra. É determinada uma rede sintáctica que é representada como uma árvore consistindo de nós ligados mutuamente por arestas e compreendendo, pelo menos, um nó superior, um ou mais nós terminais e, eventualmente, vários nós intermédios, cada nó sendo interpretado como um nó OU exclusivo servindo de apontador se existir apenas uma alternativa e servindo de ponto de escolha se existirem várias alternativas, pelo que cada um dos nós apontadores é atribuído a uma palavra da sequência e cada aresta é atribuída à relação sintáctica entre os dois nós ligados por essa aresta, pelo que cada nó é codificado por um identificador que, no caso de um nó apontador, está directamente relacionado com a entrada de uma palavra no dicionário e, no caso de um

ponto de escolha, compreende uma lista de identificadores adicionais, um dos quais tem de ser escolhido.

Do documento "*Software-Practice and Experience*", Vol. 11, 1119-1184 (1981), Donald E. Knuth e Michael F. Plass, "*Breaking Paragraphs into Lines*", é conhecido um processamento de linguagem baseado em computador, nomeadamente, composição melhorada. Este artigo discute uma nova abordagem para o problema de dividir um texto de um parágrafo em linhas de comprimento aproximadamente igual mantendo, simultaneamente, o tamanho dos espaços entre palavras dentro de uma gama desejada; evitando grandes variações no tamanho do espaçamento entre palavras encontrado numa linha em comparação com a seguinte; evitando quebras de linha "psicologicamente más" e reduzindo o número de palavras hifenizadas; evitando "pendurar" palavras em margens direitas irregulares que estão inteiramente suspensas sobre um espaço vazio na linha seguinte; e, evitando padrões feios de espaços em branco, e. g., dentados, em formato de lagarto ou de rios, que se estendem através de uma sequência de linhas e distraem o olhar do leitor. "Caixas", "cola", e "penalidades" são avaliadas com o objectivo de determinar pontos de quebra óptimos, de entre uma variedade de opções de pontos de quebra, de forma a preencher um espaço a duas dimensões tendo definido as margens esquerda e direita.

Existem várias boas razões para não colocar apenas uma frase realçada por página num livro. Estas razões podem ter sido impropriamente transferidas para a apresentação do texto em computador. A possibilidade de modificar a apresentação do texto para melhorar a aptidão do leitor em ler texto através da utilização do poder da manipulação digital aparenta ter sido ignorada pelos especialistas em computação.

É o objectivo da invenção proporcionar um método de processamento de texto para produzir, a partir de um texto legível numa máquina, um texto melhorado em que parâmetros específicos do leitor são aceites no processamento do texto, e um processador de texto para executar este método.

Para atingir este objectivo, a invenção proporciona um método de processamento de texto e um processador de texto, tal como reivindicados nas reivindicações independentes. As formas de realização preferidas da invenção estão caracterizadas nas sub-reivindicações.

É uma vantagem da invenção que segmentos de texto significativos sejam apresentados em linhas num monitor tendo em conta parâmetros específicos do leitor incluindo, pelo menos, as dimensões do campo de visualização escolhidas pelo leitor. A invenção está apta a apresentar os segmentos de texto significativos quer o campo de visualização, por exemplo a largura do monitor, seja escolhida estreita ou larga. Assim, por exemplo, podem ser apresentados segmentos de texto no écran inteiro ou numa porção mais pequena do écran em que o segmento de texto permanece significativo em termos dos segmentos de texto.

Além disso, é uma vantagem da invenção que a invenção extraia atributos de uma frase introduzida e apresente essa frase em segmentos de texto em cascata para baixo e através do écran, em que a segmentação e o deslocamento horizontal são determinados aplicando regras que utilizam partes de fala, pontuação e preferências do leitor.

As formas de realização preferidas da invenção incluem a extracção de atributos específicos do texto a partir de texto legível em computador, variando a apresentação do texto de acordo com os atributos, e criando um produto visual tridimensional (tempo e superfície de leitura) para melhorar a experiência de leitura. A cor do texto e do fundo também podem ser variadas dependendo de partes de fala e da posição das frases dentro dos parágrafos e dos parágrafos dentro de capítulos.

Nas formas de realização preferidas da invenção são utilizados atributos visuais significativos para melhorar a apresentação de texto apresentado no computador em relação à apresentação textual constante e mecânica dos sistemas actuais. É criada uma pista visual significativa relacionada com o conteúdo da nova frase e a sua relação com a frase anterior. Esta pista está disponível antes do texto ser lido, contextualizando o texto. O processamento de análise da frase utiliza a pontuação e conteúdo da frase para criar um sistema de pistas visuais significativas incluindo frases distintas que promovem um reconhecimento mais rápido das palavras numa frase e seus significados. A oração melhorada desce em cascata pela página num padrão de frases significativas determinado pelo conteúdo do texto e pelas preferências do leitor, movendo-se os olhos a curta distância de frase para frase.

Uma frase representa um pensamento completo, e um parágrafo representa um tema ou argumento discreto. A compreensão da leitura melhora quando uma frase aparece numa altura em que a transição de um parágrafo para outro é assinalada visualmente e inclui uma pausa para o aparecimento destes elementos de texto. A invenção cria atributos visuais apropriados para fornecer

estes sinais visuais. Os atributos visuais podem incluir segmentação de texto, deslocamento horizontal de uma linha relativamente a outra, cor do texto ou do fundo, brilho do texto e animação. A animação pode incluir cintilação e a iluminação progressiva, dependente do tempo e temporária, de elementos do texto, e o movimento do texto de uma apresentação padrão para uma em cascata.

Numa sessão de leitura exemplificativa, o leitor selecciona o texto a ler, e informa o sistema de leitura do tipo de texto, e. g., um romance. O sistema de leitura recupera o ambiente apropriado previamente armazenado para o leitor que está a ler o tipo de texto seleccionado. O leitor tem a oportunidade de editar as regras de leitura e conjuntos de palavras mas recusa. Começando no início de um capítulo, onde o leitor tinha parado, o texto é apresentado, uma frase de cada vez, em cascata para baixo e através do écran. Como as linhas apresentadas são partidas em segmentos de texto significativos em vez de terminar arbitrariamente em 80 caracteres, é possível ler e compreender um segmento completo de uma só vez, em que o movimento dos olhos resultante é feito desde o centro de segmento de texto ao centro de segmento de texto, para baixo e através do écran. A cor de fundo da frase é uma função da posição da frase dentro do parágrafo e da posição do parágrafo dentro do capítulo. Cada uma das frases num parágrafo pode ter a mesma tonalidade de cor, mas ter diferenças incrementais identificáveis na saturação da cor ou escuridão entre frases, e diferenças incrementais identificáveis na tonalidade de cor entre parágrafos. A cor de fundo proporciona, assim, informação posicional não literal, assinalando o início e o fim de um parágrafo, e o início e fim de um capítulo.

As frases são segmentadas de acordo com as regras aprovadas pelo leitor. Cada preposição dobra uma frase de uma forma previsível, resultando em dois segmentos de texto, cada um situado numa nova linha. O conteúdo da linha de cima tem efeito no deslocamento horizontal da linha de baixo. A mesma preposição dobra e desloca horizontalmente texto da mesma maneira sempre que possível. Como resultado, a mesma preposição provoca um padrão semelhante, há muito familiar ao leitor. Partes de fala têm efeito na cor do texto de acordo com regras aprovadas pelo leitor. Como resultado, tanto a estrutura da frase como partes de fala são imediatamente reconhecíveis a partir do reconhecimento de padrões e cores, mesmo antes de uma palavra do texto ser reconhecida, fornecendo pistas visuais não literais ao leitor.

As frases avançam a um determinado ritmo utilizando uma regra previamente seleccionada pelo leitor. A fórmula é uma função do tipo de texto, número de palavras, nível educacional, e número de sílabas presentes na linha. O ritmo de avanço é um pouco mais rápido do que o ritmo de conforto do leitor, para aumentar a compreensão. Uma pausa mais longa entre a apresentação de duas frases assinala um novo parágrafo antes de ele aparecer. O leitor é capaz de interagir facilmente com o sistema de leitura, mantendo frases difíceis no écran durante mais tempo, e acelerando ou travando a apresentação.

A sessão de leitura exemplificativa descrita acima é exemplificativa, não limitando ou esgotando as possíveis formas de realização dentro do âmbito da invenção.

O processo de melhoria explora o facto de que os elementos de texto têm qualidades não literais que podem ser utilizadas

para produzir um produto de apresentação de texto no tempo e no espaço que seja mais significativo e que melhore a aptidão do leitor para compreender o significado literal do texto a um nível superior ao de formatos existentes em computadores ou papel. A cascata de frases para baixo e através do écran confere um corolário visual de como o texto poderia ser recebido se fosse ouvido numa leitura ao vivo, sem o leitor ter de o sussurrar. A leitura de um texto assim melhorado é quase como ouvir com os olhos, mas como um pacote visual total que é experimentado quase simultaneamente.

As formas de realização preferidas da invenção são agora descritas fazendo referência aos desenhos nos quais:

A FIG. 1 é um diagrama de fluxo de alto nível de um método utilizado numa forma de realização da invenção;

A FIG. 2 é uma tabela que ilustra vários atributos dos elementos de uma frase;

A FIG. 3 contém um exemplo de texto melhorado incluindo um cascata e ângulos descendentes para a direita;

A FIG. 4 apresenta a frase da FIG. 3 como texto melhorado incluindo uma cascata e ângulos descendentes para a direita e para a esquerda;

A FIG. 5 ilustra o passo 110 em detalhe;

A FIG. 6 ilustra conjuntos de palavras que estabelecem uma hierarquia de pontos de dobragem secundários;

A FIG. 7 ilustra as localizações dos pontos de dobragem numa frase;

A FIG. 8 ilustra o passo 136 em detalhe;

A FIG. 9 ilustra a frase da FIG.7 apresentada como texto melhorado;

A FIG. 10 ilustra um Mapa Histórico de Amostragem do Intervalo de Apresentação, em que o tempo necessário para ler é uma função da dificuldade e complexidade da frase; e

A FIG. 11 ilustra a frase da FIG.7 mostrando 5 pronúncias visuais de sílabas enfatizadas.

A FIG. 1 mostra um diagrama de fluxo de alto nível de um método utilizado numa forma de realização da invenção. Os passos da FIG. 1 e todas as figuras deste tipo podem, é claro, ser repartidas em passos adicionais ou fundidas em menos passos e ainda assim descreverem a mesma forma de realização.

No passo 102, o dispositivo de leitura é iniciado, e é passada a identificação (ID) do leitor e a identificação do ficheiro de texto a ser lido. Numa forma de realização preferida, a ID do leitor é obtida a partir do sistema operativo. Noutra forma de realização o leitor é incitado a identificar-se. Numa forma de realização, existe apenas um leitor do dispositivo assumido. Dada a ID do leitor, os parâmetros específicos do leitor são extraídos no passo 104. Os parâmetros são discutidos em detalhe adiante. Dada a ID do texto, o ficheiro de texto e os parâmetros específicos do texto são extraídos no passo 104. Numa forma de realização preferida, os

parâmetros têm valores predeterminados, e não requerem entrada de informação pelo leitor. Numa forma de realização muito preferida, a entrada pelo leitor, de parâmetros específicos preferidos é permitida, os valores aceites, e os valores dos parâmetros armazenados tal como no passo 108. Numa forma de realização, o utilizador pode activar menus contendo caixas de diálogo, ver os parâmetros predeterminados, e modificá-los utilizando interfaces de utilizador normalizadas incluindo entradas numéricas, entradas de texto, barras de ferramentas, e botões. Numa forma de realização preferida, dada uma permissão de acesso adequada, é também permitida a modificação de parâmetros globais que se aplicam a todos os leitores do texto. Ao leitor é permitido ter parâmetros estabelecidos apenas para o texto actual, para textos do mesmo tipo que o texto actual, ou para todos os textos lidos por este leitor.

Os parâmetros específicos do leitor incluem: as dimensões do campo de visualização; a paleta de cores para o fundo e o texto; as entradas de conjuntos de palavras para as regras; o comprimento mínimo da frase; o comprimento máximo de linha; o peso mínimo e máximo de um segmento de texto; os ângulos de descida; as entradas das regras de deslocamento horizontal; as preferências de justificação horizontal; o comprimento do intervalo entre frases; o ritmo de avanço; as proporções dos intervalos entre frases e entre parágrafos; as definições de etiquetagem; as especificações de animação; as definições de parâmetros para derivar os intervalos da apresentação dependentes do conteúdo; as definições de parâmetros e de pesos para derivar o assunto, o peso, a densidade, a complexidade, e o conteúdo da frase/texto; e as identidades de textos de acontecimentos especiais.

Dimensões do Campo de Visualização

As dimensões do campo de visualização podem especificar o tamanho da janela para visualização do texto nas duas dimensões X e Y e numa terceira dimensão, o tempo. Este campo pode incluir o écran inteiro ou uma porção mais pequena do écran. A dimensão tempo pode incluir o tempo durante o qual ocorrem animações dependentes do tempo tal como o piscar.

Paleta de Cores

A paleta de cores para o fundo e o texto permite a especificação da cor preferida do fundo e do texto que varia com o texto. Numa forma de realização preferida, a tonalidade de cor e o brilho da cor do fundo variam com a posição de uma frase dentro de um parágrafo e com a posição do parágrafo dentro de um capítulo. Numa forma de realização muito preferida, o brilho é colocado num nível para a primeira frase num parágrafo e num segundo nível para a última frase do parágrafo, com o brilho de cada frase intermédia progredindo do primeiro nível até ao segundo nível. Numa forma de realização preferida, a tonalidade de cor do fundo é colocada num primeiro valor para o primeiro parágrafo num capítulo e num segundo valor para o último parágrafo num capítulo, com a tonalidade de cor de cada parágrafo progredindo do primeiro valor até ao segundo valor. Numa forma de realização muito preferida, a primeira tonalidade de cor é verde e a segunda tonalidade de cor é violeta.

Em formas de realização preferidas, a cor do texto pode variar com os atributos categóricos e contínuos da palavra. Um

atributo categórico é uma função da categoria a que uma palavra pertence. Numa forma de realização muito preferida, as categorias incluem partes de fala e. g. verbos, e palavras técnicas. Os atributos contínuos incluem aqueles que podem ser medidos, tal como o nível educacional e. g. décimo segundo ano, número de sílabas, e comprimento da palavra. Numa forma de realização preferida, diferentes partes de fala têm diferentes cores de texto. Numa forma de realização muito preferida, a cor do texto é variada juntamente com a cor do fundo para proporcionar um contraste óptimo preferido pelo leitor.

Podem ser utilizados conjuntos de palavras para especificar algumas partes de fala. Por exemplo, "I" é um pronome, "had" é um verbo, "a" é um artigo, e "farm" pode ser um substantivo, verbo, ou adjectivo. As partes de fala necessitam apenas de ser partes prováveis de fala, tal como no caso da palavra "farm", que pode ser um substantivo, verbo, ou adjectivo. Na FIG. 2 é mostrado um exemplo de partes de fala numa frase única. São mostradas as múltiplas partes possíveis da coluna 150 de fala, bem como a parte não ambígua da coluna 152 de fala.

São introduzidos conjuntos de palavras são introduzidas em formas de realização preferidas, utilizando dois métodos. O primeiro método utiliza um formato de texto estruturado, tal como a *Standard Generalized Mark-up Language* (SGML), permitindo a importação e exportação de conjuntos grandes de palavras e dicionários de fontes diversas. Os analisadores gramaticais padrão SGML são então utilizados para armazenar os conjuntos de palavras numa base de dados apropriada. O segundo método permite a edição interactiva dos conjuntos de palavras utilizando uma hierarquia de menus e caixas de diálogo. Numa forma de realização preferida, os conjuntos de palavras especificados

pelo leitor são armazenados numa base de dados separada dos conjuntos de palavras do dicionário ou glossário, para permitir o armazenamento, edição e recuperação mais rápida em separado.

Comprimento Mínimo de um Segmento de Texto e Comprimento Máximo de Linha

Uma forma de realização preferida inclui um comprimento mínimo do segmento de texto especificado pelo leitor. Um segmento de texto é um pedaço de texto que é apresentado numa linha num monitor. O comprimento mínimo do segmento de texto, que pode ser denominado em palavras, é o mínimo que o leitor deseja ver numa linha separada. Um comprimento máximo de linha é o comprimento máximo que o leitor deseja ver numa linha. Os comprimentos podem ser medidos em número de caracteres, palavras, sílabas, nível educacional, ou qualquer soma de produtos destes. Numa forma de realização, o comprimento mínimo de linha é dez caracteres e o comprimento máximo de linha é trinta e cinco caracteres.

Peso da Frase

O peso da frase é um atributo derivado de uma frase (segmento de texto ou segmento de texto potencial) fornecendo alguma medida da quantidade de material numa frase. Numa forma de realização, o peso da frase é simplesmente o número de palavras numa frase. Numa forma de realização preferida, o peso da frase inclui a densidade da frase e a complexidade da frase. A densidade da frase pode incluir o número de palavras técnicas ou o número de palavras excedendo um certo nível de

qualificação. A complexidade da frase pode incluir o número de semelhanças de ortografia entre palavras numa frase, o número de palavras ambíguas, e o peso total de palavras com o peso especificado pelo leitor.

Peso Mínimo e Máximo de um Segmento de Texto

Uma forma de realização preferida inclui um peso mínimo e máximo de um segmento de texto especificado pelo leitor. O peso do segmento de texto é a quantidade de texto que o leitor deseja ver numa linha separada. Em formas de realização preferidas, o peso do segmento de texto é a soma de pesos das frases dentro do segmento de texto.

Regras de Justificação Horizontal

As regras de justificação horizontal especificam a justificação horizontal de uma linha de texto relativamente à linha acima. A justificação pode incluir o tipo de justificação para a linha ou a frase a ser posicionada, *i. e.* justificação à direita, à esquerda ou ao centro. A justificação também pode incluir a porção do texto da linha acima a partir da qual a justificação é medida, *i. e.* a linha inteira de texto ou uma frase, versus o ponto dessa porção utilizada, *e. g.*, o mais à esquerda, o mais à direita, ou centro. A justificação horizontal numa forma de realização é simplesmente medida dentro da linha a ser posicionada em vez de relativamente à linha acima.

Numa forma de realização preferida, a primeira frase numa linha é justificada ao centro, medida a partir do centro da

última frase na linha imediatamente acima. Noutra forma de realização, a linha inteira de texto é justificada ao centro por baixo do centro da linha de texto acima. Ainda noutra forma de realização, o segmento de texto "centro de gravidade", calculado utilizando a dificuldade de cada palavra, é utilizado como centro do segmento de texto para fins de justificação.

Ângulo de Descida

Um ângulo de descida é a quantidade de deslocamento horizontal para cada nova linha, modificando a posição horizontal exigida somente pelas regras de justificação horizontal. Por definição, cada segmento de texto é apresentado numa nova linha. Numa forma de realização preferida, o ângulo de descida é especificado em unidades de caracteres. O ângulo de descida e a justificação horizontal determinam, pelo menos parcialmente, "a cascata de texto" para baixo e através do écran em formas de realização preferidas. Nas FIGS. 3 e 4 é ilustrado um exemplo de texto em cascata. Um ângulo de descida pode ser zero, significando que, sem mais, a posição horizontal do segmento de texto é determinada apenas pelas regras de justificação horizontal. Um ângulo de descida pode ser esquerdo, em que a linha abaixo é desviada para a esquerda relativamente à linha acima, ou direito, em que o texto é desviado para a direita. A FIG. 3 ilustra apenas ângulos de descida direitos.

Numa forma de realização, o ângulo de descida é uma constante para cada nova linha. Numa forma de realização preferida, o ângulo de descida é uma função do peso do segmento de texto da linha acima. Noutra forma de realização preferida, as regras de justificação horizontal exigem justificação ao

centro por baixo do centro de cada linha imediatamente acima, e o ângulo de descida é calculado para apresentar um caminho substancialmente recto, quando todas as linhas de texto são apresentadas, do centro de linha para o centro de linha, do canto superior esquerdo para o canto inferior direito na superfície do monitor.

Numa forma de realização preferida, as entradas para as regras do ângulo de descida incluem atributos do texto na linha acima. Numa forma de realização preferida, as entradas incluem a razão para dobragemr a linha acima, *i. e.* o ponto de dobragem primário, o ponto de dobragem secundário, ou a regra de colapso. Numa forma de realização preferida, é exigido um ângulo de descida mais positivo quando a linha imediatamente acima é dobragemda devido a um ponto de dobragem primário do que quando é dobragemda devido a um ponto de dobragem secundário. Noutra forma de realização preferida, as entradas incluem o peso do segmento de texto da linha actual e da linha acima.

É reconhecido qua a regra de justificação horizontal podia exigir a justificação à esquerda e a medição do deslocamento horizontal a partir da margem esquerda, bem como um ângulo de descida zero, combinando-se para resultar em texto justificado à esquerda em cada linha. É também reconhecido que o posicionamento do texto na horizontal pode ser conseguido de várias maneiras equivalentes às do exemplo acima. Em particular, os cálculos da posição do texto podem ser conseguidos justificando primeiro e depois desviando, ou desviando primeiro e depois justificando com resultados equivalentes.

Comprimento do Intervalo

Numa forma de realização, os intervalos estão associados a pontos de dobragem cujas posições foram determinadas, mas que devido a outras regras, permanecem na mesma linha e não provocam a dobragem. Um intervalo de zero ou mais espaços é adicionado após um ponto de dobragem, onde esse ponto de dobragem não provocou a criação de uma nova linha. Numa forma de realização preferida, o comprimento do intervalo é um parâmetro determinado pelo leitor, em que um comprimento de intervalo zero não resulta na criação de intervalos. Os intervalos permitem uma pista visual quanto à existência de frases, mesmo quando as frases não causaram a formação de uma nova linha.

Ritmo de Avanço

Os ritmos de avanço especificam os tempos de duração de apresentação e os intervalos de tempo entre a apresentação de um agrupamento de texto e o seguinte. Numa forma de realização preferida, uma frase é apresentada num écran. Numa forma de realização preferida, o ritmo de duração da apresentação pode ser uma função do número de palavras do texto apropriadamente pesadas, do nível educacional, do tempo de pronúncia, do número de frases, do número de sílabas, ou do peso da frase. Os intervalos de tempo podem incluir o intervalo entre frases e parágrafos. Numa forma de realização preferida, o intervalo de tempo entre frases e parágrafos é diferente. Desta forma, o ritmo constante de chegada de texto nos sistemas actuais pode ser substituído por um ritmo dependente de onde se encontra o texto e qual é o conteúdo do texto.

Estações

As estações são posições na superfície do monitor nas quais o texto é apresentado. As estações podem ser janelas identificadas dentro das quais é apresentado texto ou pontos nos quais começa a apresentação de texto. Uma forma de realização preferida inclui uma estação de texto activa, onde é apresentado o texto a ser lido. Outra forma de realização preferida inclui uma estação de pré-leitura apresentando texto prestes a ser lido e uma estação de pós-leitura apresentando texto já lido. Numa forma de realização, o texto total apresentado em todas as estações na superfície de visualização é controlado pelo tempo esperado de leitura ou de pronúncia. Numa forma de realização, apenas cerca de um minuto de material é permitido no écran de uma só vez.

Animação

A animação é a apresentação de texto dependente do tempo. Exemplos de animação incluem cintilação, dissolução, e o apagamento, isto é, a iluminação de porções de texto sucessivas a um ritmo especificado. Numa forma de realização preferida, um efeito de deslizar modificado é conseguido empilhando texto prestes a ser lido em linhas numa estação de pré-leitura no topo do écran, empilhando as linhas previamente lidas numa estação de pós-leitura no fundo do écran, e colocando em cascata a frase activa numa estação activa que ocupa o meio do écran. Numa forma de realização, as cores do texto de pré-leitura e pós-leitura e/ou do fundo diferem daquela do texto que está a ser lido presentemente.

Etiquetagem

A etiquetagem inclui designar se uma frase é etiquetada com uma definição ou desenho relacionado com a palavra. Numa forma de realização preferida, um conjunto de palavras especifica que palavras serão etiquetadas com uma definição. Quando uma palavra etiquetada é seleccionada por um leitor, pode ser apresentada uma definição, desenho ou outra ilustração electrónica da palavra. Numa forma de realização preferida, uma palavra etiquetada, quando seleccionada, apresenta uma janela contendo a definição ou desenho. Os métodos de selecção preferidos incluem a utilização de um rato e do botão *shift* da direita.

Acontecimento de Leitura

Um acontecimento de leitura inclui o aparecimento e desaparecimento de uma frase e qualquer tempo adicional antes do aparecimento da frase seguinte. Um acontecimento de leitura contém uma série de momentos de leitura. Um momento de leitura é um aparecimento de texto estático para os olhos do leitor que tenta compreender o texto. Para muitos leitores, o momento de leitura encontrará os olhos centrados numa frase, a frase sozinha numa linha, o leitor compreendendo a única frase.

Ritmo de Leitura

O intervalo de apresentação é a duração da apresentação de uma frase no écran. O intervalo de apresentação pode ser

controlado pelo leitor ou determinado automaticamente. O controlo pelo leitor, numa forma de realização, pode ser feito através de toques no rato. Noutra forma de realização, o controlo pelo leitor pode ser através da detecção do movimento dos olhos ou da detecção de mudanças nas ondas cerebrais. A detecção de ondas cerebrais opera gerando um pedido de nova frase quando as ondas cerebrais correspondendo à actividade de leitura mudam para ondas cerebrais associadas com ter acabado de ler a frase.

O ritmo de leitura é medido em conteúdo de texto por unidade de tempo. Estas Unidades de Conteúdo de Texto podem ser medidas em unidades de comprimento, tais como o número de palavras, sílabas, tempo de pronúncia, ou qualquer das medidas previamente discutidas de comprimento ou peso da frase ou qualquer soma de produtos destes. As Unidades de Conteúdo de Texto também podem ser medidas em unidades de complexidade, tais como as discutidas previamente em relação à complexidade ou peso da frase ou qualquer soma de produtos destas. As Unidades de Conteúdo de Texto numa forma de realização preferida são calculadas para uma frase apresentada e utilizadas para comparar o verdadeiro ritmo de leitura em Unidades de Conteúdo de Texto por tempo com intervalos de apresentação controlados pelo leitor, e colocar o intervalo de apresentação do texto para intervalos de apresentação do texto controlados automaticamente. Desta forma, o intervalo de apresentação pode ser tornado dependente do conteúdo, e pode ser feito seguir o ritmo provável de apresentação preferido do leitor.

Acontecimentos Especiais

Os acontecimentos especiais são acontecimentos dependentes do tempo que ocorrem durante o acontecimento de leitura. Os acontecimentos especiais incluem uma pista visual ou audível dando um sinal que permanece no tempo. Por exemplo, quando tiver decorrido 75% do intervalo de apresentação dependente do conteúdo, aparece uma pista visual no écran. Os acontecimentos especiais podem ser organizados por um guião de acontecimentos especiais. Os acontecimentos especiais complexos incluem uma série de acontecimentos especiais.

Pronúncia Visual

A pronúncia visual envolve variar no tempo as mudanças de cor ou brilho de frases, palavras, e sílabas individuais, apesar de não necessariamente todas. Uma forma de realização dá ênfase a palavras mais compridas e mais difíceis que requerem mais tempo para serem pronunciadas utilizando o sublinhado. A pronúncia visual é um exemplo de um acontecimento especial complexo.

Puxar a Cascata para baixo

Numa forma de realização, o texto é apresentado numa estação de pré-leitura para uma percentagem do intervalo de apresentação dependente do conteúdo, e depois é apresentado numa cascata de frases com um segmento de texto ou uma linha de cada vez. Puxar a cascata para baixo é um exemplo de um acontecimento especial complexo.

Texto Não-Linear

O texto linear é um texto apresentado e lido do início ao fim. Um livro destinado a ser lido do início ao fim é um exemplo comum de um texto linear. O texto não-linear inclui a apresentação e leitura de um texto numa ordem diferente da linear. Um exemplo de um texto não-linear é o Hipertexto, em que certas palavras são apresentadas de forma a indicar ligação e. g. o texto é inclinado ou de cor azul. A apresentação de texto não-linear pode ser dirigida pelo leitor ou automática. Um exemplo disto é a apresentação automática de Hipertexto numa profundidade de primeira ordem, seguindo a primeira ligação de Hipertexto, ignorando os ciclos. Outro exemplo é uma primeira apresentação alargada, apresentando o primeiro nível de texto, seguido das ligações de Hipertexto do nível principal, seguido de quaisquer ligações de Hipertexto do primeiro nível, e continuando até ser atingida a profundidade mais baixa ou até o leitor intervir. Ainda outro exemplo é a depuração da profundidade ou do alargamento da primeira apresentação para incluir apenas porções tendo certas palavras-chave de interesse.

Lupa Mágica

Numa forma de realização, a apresentação de texto pode começar pela apresentação de texto comum não formatado. Quando o leitor deseja ver o texto apresentado como texto melhorado, o leitor designa uma porção de texto para melhoramento colocando um ícone, tal como um "ícone lupa mágica", sobre alguma parte da frase. Essa frase e as frases seguintes são apresentadas como

texto melhorado. Esta apresentação de texto melhorado continua até o leitor desactivar a lupa mágica. Esta lupa mágica pode servir como um ponto de entrada alternativo no processador de texto melhorado.

Método

Na FIG. 1, no passo 110, o texto é pré-processado. O passo 110 é mostrado em detalhe na FIG. 5. O texto é descrito para identificar parágrafos, frases, palavras, e pontuação. Os parágrafos podem ser identificados por linhas em branco, marcadores de parágrafos, caracteres de indentação, caracteres de tabulação, ou qualquer outra característica adequada no texto. As frases podem ser identificadas utilizando regras gramaticais incluindo períodos, espaçamento, primeiras palavras maiúsculas, e abreviações ou ausência destas. Numa forma de realização preferida a leitura de texto corrente, período, ponto de interrogação ou ponto de exclamação, sejam isolados ou seguidos por um período, seguido de dois espaços ou de fim de parágrafo, assinala o fim de uma frase.

No passo 124, a frase é simbolizada em palavras e pontuação. A ênfase especificada pelo autor original, e. g. itálico ou sublinhado, é preservada nas formas de realização preferidas. Numa forma de realização é utilizado um amostrador lexical padrão tal como o Lex (MARCA REGISTADA), em que o fim de uma palavra é indicado nas regras gramaticais por um espaço em branco ou pontuação. Outra forma de realização utiliza um analisador lexical manuscrito. Uma forma de realização armazena caracteres de formatação, tais como tabulações ou indentações, como pontuação. A posição de uma palavra é armazenada, de um

modo preferido, como um atributo da palavra, para proporcionar ligações para, e fazer uma busca no trabalho original.

Uma forma de realização preferida também permite que grupos de palavras sejam “presos” conjuntamente, e sejam reconhecidos como um grupo de palavras. Numa forma de realização, tais grupos de palavras são reconhecidos pelo amostrador lexical. Noutra forma de realização, tais palavras são reconhecidas por um pré-processador precedendo o amostrador lexical para garantir o reconhecimento como frase em vez de meramente como palavras individuais. As palavras presas, por exemplo, “Príncipe de Gales”, seriam reconhecidas como uma frase única, e, de um modo preferido, não são partidas pela preposição em duas frases apresentadas em duas linhas.

No passo 126, as palavras são procuradas em dicionários, glossários e tabelas para determinar os atributos da palavra. O texto é além disso processado para determinar os atributos categóricos e contínuos. Numa forma de realização preferida, os atributos categóricos importantes incluem partes de fala, e os atributos contínuos importantes incluem a posição da palavra, o nível educacional, o tempo de pronúncia, e o número, a posição e o som das sílabas, e o nível de ênfase vocal. Identificar partes de fala com 100% de precisão pode exigir programação intensiva para determinar o contexto do mundo real do texto. Esta precisão não é exigida para a prática da invenção, uma vez que os erros são de consequência menor porque o leitor é um ser humano, não uma máquina. As partes possíveis de fala são determinadas primeiro procurando a palavra num dicionário ou glossário. Este dicionário ou glossário necessita apenas de possuir partes prováveis de fala para uma palavra, e não uma definição. Por exemplo, a palavra “farm” pode ser um substantivo, verbo ou

adjectivo. Uma forma de realização preferida armazena as partes dos atributos da fala utilizando um mapa de *bits* para preservar as múltiplas partes possíveis de fala. Uma forma de realização armazena explicitamente um atributo de ambiguidade, indicando se a palavra ainda tem múltiplas partes possíveis de fala. Outra forma de realização utiliza a existência de mais do que uma única parte possível de fala como uma indicação de ambiguidade.

Numa forma de realização preferida, partes de fala predeterminadas existem num dicionário e podem ser procuradas. Numa forma de realização muito preferida, um conjunto de palavras pode ser adicionado para se sobrepor ou completar o conjunto predeterminado. Noutra forma de realização, as palavras técnicas são especificadas por conjuntos de palavras introduzidas pelo utilizador.

A FIG. 6 ilustra nove conjuntos de palavras especificando partes de fala, organizadas numa hierarquia. Estes conjuntos de palavras e hierarquias também são utilizados como entrada para as regras de dobragem, descritas mais tarde. Uma forma de realização preferida verifica os conjuntos de palavras ilustrados na FIG. 6, começando na Classe 1, Subclasse 1, e acabando com a Classe 3, Subclasse 1. A busca é terminada quando uma palavra ou frase é encontrada. Dado o vocabulário limitado e a natureza estácticos conjuntos de palavras na FIG. 6, uma forma de realização muito preferida utiliza um analisador gramatical manuscrito rápido para pesquisar pelas palavras e frases na FIG. 6. AS formas de realização preferidas incluem analisadores gramaticais escritos em C, C++, Perl, AWK compilado, AWK e AWK a C, C com funções de expressão regulares, ou qualquer linguagem adequada. Uma forma de realização de um analisador gramatical utiliza a YACC.

Numa forma de realização, o dicionário é um dicionário disponível comercialmente em meio electrónico, tal como o CD-ROM. O dicionário padrão é analisado quanto aos atributos de palavra tais como partes de fala e número de sílabas. Como as definições das palavras não são necessárias em muitas formas de realização, é possível o armazenamento de várias palavras com o número de sílabas e partes de fala associados. Numa forma de realização muito preferida, as palavras geralmente mais utilizadas e as mais recentemente utilizadas são armazenadas numa memória de acesso rápido, tal como uma memória de acesso aleatório (RAM) de estado sólido. Em formas de realização nas quais os dicionários serão manuscritos, é preferido um método rápido utilizando endereçamento calculado (*hashing*), detecção de colisão e estruturas de armazenamento (*buckets*). Em formas de realização nas quais os conjuntos de palavras são fixados antes da leitura, é preferido o endereçamento calculado perfeito sem a utilização de estruturas de armazenamento.

Ainda noutra forma de realização, o nível de ênfase na pronúncia é derivado como um atributo dependendo, em parte, da parte de fala. Numa forma de realização muito preferida, a ênfase de pronúncia é categorizada como primária, secundária, e nenhuma. Numa forma de realização, o tempo de pronúncia e o som real, e. g. tal como encontrado num ficheiro de som, são também extraídos do dicionário ou glossário e armazenados como atributos da palavra.

No passo 128, são pesquisados os conjuntos de palavras específicos do texto e do leitor. Numa forma de realização preferida, os conjuntos de palavras especificados pelo leitor são armazenados numa base de dados separada dos conjuntos de

palavras do dicionário ou glossário, para permitir o armazenamento separado e recuperação mais rápida. Numa forma de realização, os conjuntos de palavras especificados pelo leitor são verificados antes dos dicionários, e os dicionários apenas verificados se palavras e atributos necessários não forem encontrados nos conjuntos de palavras especificados pelo leitor. As formas de realização preferidas utilizam uma hierarquia de base de dados para procurar palavras.

No passo 130, múltiplas partes de fala são tornadas não ambíguas. Numa forma de realização, é utilizada uma microgramática para determinar as partes de fala prováveis. Uma microgramática utiliza palavras adjacentes ou próximas para determinar com mais precisão a parte de fala mais provável para uma palavra. Por exemplo, a palavra "*farm*" na frase "*a farm in*" seria provavelmente um substantivo uma vez que é precedida por um artigo e sucedida por uma preposição. Como outro exemplo, se uma palavra pode ser ou um substantivo ou um verbo, e a palavra for precedida por "*could*", "*will*", "*shall*", ou "*to*", então a palavra é provavelmente um verbo. Se a palavra "*farm*" fosse precedida por "*will*", a palavra era provavelmente um verbo. Noutra forma de realização, toda a não ambiguidade é feita simplesmente escolhendo a utilização da palavra estatisticamente mais provável. Em ainda outra forma de realização, não existe não ambiguidade automática, apenas não ambiguidade manual utilizando edição humana. Numa forma de realização preferida, é armazenado um atributo de ambiguidade para cada palavra, indicando se ainda existem múltiplas partes de fala possíveis após a não ambiguidade. Em ainda outra forma de realização, um atributo de ambiguidade não é armazenado mas derivado da existência de múltiplas partes de fala possíveis armazenadas para uma palavra. Numa forma de realização, a ambiguidade é

inferida a partir da apresentação visual de cores de texto listadas ou alternadas associadas com cada parte de fala. Por exemplo, se os verbos forem laranja e os adjetivos forem amarelos, então um possível verbo ou adjetivo pode ter listas ou caracteres de texto amarelos e laranja alternados.

Numa forma de realização da invenção, partes de fala são determinadas em parte pesquisando morfemas (palavras de raiz) e atribuindo parte dos atributos de fala com base nas terminações e. g. *-ly*, *-ing*.

No passo 132, os atributos determinados para as palavras e frases são armazenados, criando uma "frase enriquecida", que provavelmente, permanecerá inalterada entre leitores. Por exemplo, o nível educacional e parte de fala de uma palavra permanecerão inalterados para diferentes leitores, apesar de poderem desejar comprimentos de segmento de texto e ritmos de apresentação diferentes. Numa forma de realização, a frase enriquecida é armazenada numa memória persistente tal como um ficheiro. Noutra forma de realização, o texto enriquecido é armazenado num CD-ROM. Numa forma de realização preferida, a frase enriquecida é implementada como uma lista ligada de nós, tendo cada nó os atributos de palavra e frase descritos acima incluindo a posição da palavra no texto original. Na FIG. 5, passo 134, os pontos de dobragem primários são determinados aplicando regras de pontos de dobragem primários. Os pontos de dobragem são pontos divisores de texto localizados entre letras. Numa forma de realização preferida, os pontos de dobragem são classificados como primários e secundários. Os pontos de dobragem primários são determinados utilizando as regras de dobragem primárias que determinam as posições dos pontos de dobragem primários com base em marcas de pontuação. A FIG. 7

ilustra um ponto de dobragem primário depois da vírgula a seguir a "Africa". Os pontos de dobragem primários dividem o texto em "Super frases". Numa forma de realização preferida, os pontos de dobragem primários estão situados em cada vírgula, dois pontos, ponto e vírgula, e parêntesis, parêntesis recto ou chaveta esquerda. A posição do ponto de dobragem pode ser armazenada como um atributo num nó numa lista ligada de nós formando uma frase enriquecida.

Os pontos de dobragem secundários são determinados aplicando regras de pontos de dobragem secundários. Em formas de realização preferidas, os pontos de dobragem e as regras secundárias são ordenados numa hierarquia e as regras de dobragem secundárias aceitam partes de fala como entradas. Numa forma de realização muito preferida, as regras de dobragem secundárias incluem como regra atributos de entrada do conteúdo do texto, dos segmentos do texto e das frases a serem processadas. Por exemplo, um ponto de dobragem secundário pode ser exigido por um segmento de texto excedendo o peso máximo de um segmento de texto preferido pelo leitor apesar de um comprimento máximo de segmento de texto não ter sido atingido.

Os atributos contínuos, tais como a dificuldade, densidade, complexidade, potência e tempo de pronúncia da frase podem ser utilizados como entradas para uma regra modificando a ordenação estabelecida por uma tabela, tal como a da FIG. 6, utilizando apenas partes de fala para determinar a ordenação das partes de dobragem secundárias. Por exemplo, um segmento de texto tendo um peso maior do que 35 por cento acima da média do texto teria uma ordem de Classe 1 atribuída independentemente da ordenação de outra forma exigida pela tabela da FIG. 6. Numa forma de realização preferida, o peso ou a potência da frase é utilizada

exclusivamente para determinar a ordenação de partes de dobragem secundárias, em vez de unicamente partes de fala.

Numa forma de realização alternativa, as regras de dobragem exigem uma dobragem baseada no número de caracteres na linha, e as partes de fala são apresentadas utilizando cores correspondentes a uma parte da fala da palavra. A última forma de realização pode não oferecer as vantagens da cascata, mas oferece pistas de apresentação visuais baseadas no conteúdo do texto.

A FIG. 6 ilustra uma tabela utilizada numa forma de realização preferida para determinar os pontos de dobragem secundários. Por exemplo, as preposições determinam os pontos de dobragem secundários de Classe 3, Subclasse 1. Na FIG. 7, existem pontos de dobragem secundários de classe 3 antes das proposições "in" e "of". Os pontos de dobragem secundários dividem as Super-frases em "Mini-frases", como ilustrado na FIG. 7. As Mini-frases estão relacionadas com segmentos de texto em que as Mini-frases são muitas vezes o mesmo que os segmentos de texto e são muitas vezes apresentadas, cada uma, numa nova linha. Contudo, depois das Mini-frases serem identificadas, as regras podem ditar o aparecimento de mais ou menos uma Mini-frase como um segmento de texto numa linha.

As regras de dobragem primárias são aplicadas primeiro, seguidas pelas regras de dobragem secundárias, aplicadas por ordem da ordenação de regras de dobragem. Na FIG. 6 é mostrado um exemplo de ordenação de regras de dobragem secundárias, estabelecido pela ordenação dos conjuntos de palavras dando origem aos pontos de dobragem secundários. Algumas formas de realização preferidas utilizam ou o peso ou a potência da frase

para determinar a ordenação dos pontos de dobragem secundários em vez de utilizar unicamente partes de fala. Uma forma de realização muito preferida permite a entrada de uma preferência do leitor para a determinação, da ordenação dos pontos de dobragem secundários por partes de fala ou peso/potência da frase. Alguns leitores preferem a segmentação do texto baseada na estrutura, enquanto outros preferem a segmentação baseada na complexidade ou no tempo estimado para ler um segmento de texto.

Numa forma de realização preferida, as regras de dobragem secundárias são aplicadas apenas até ser alcançado um limite. Este limite é muitas vezes o comprimento de linha mínimo. Na FIG. 5, passo 136, e em detalhe na FIG. 8, passo 136 é mostrado um método para determinar a posição dos pontos de dobragem secundários. Numa forma de realização em que a aplicação de uma regra de dobragem secundária a uma Super-frase resultaria num comprimento de Mini-frase menor que o comprimento de linha mínimo especificado, a regra de dobragem não é aplicada e não são aplicadas regras de dobragem adicionais a essa Super-frase. Pelo contrário, quando não existir de outra forma um ponto de dobragem numa linha excedendo o comprimento de linha máximo, é aplicada uma regra de colapso, forçando a dobragem do texto em duas linhas. Quando todas as Super-frases não tiverem regras de dobragem adicionais aplicadas, o processo de dobragem está completo.

Na FIG. 1, passo 111, os parâmetros, atributos, e regras de dobragem podem ser utilizados como entrada para as regras de deslocamento horizontal. As regras de deslocamento horizontal determinam a posição horizontal do segmento de texto. Numa forma de realização preferida, as regras de deslocamento horizontal incluem tanto as regras de justificação horizontal como as

regras de ângulo de descida. O deslocamento horizontal nesta forma de realização é a soma dos resultados da regra de justificação horizontal com a regra do ângulo de descida. O deslocamento horizontal nesta forma de realização é a soma dos resultados da regra de justificação horizontal com a regra do ângulo de descida. Numa forma de realização fácil de implementar, a regra de deslocamento horizontal é simplesmente o ângulo de descida tal como aplicado ao segmento de texto justificado ao centro. Esta forma de realização não utiliza a regra de dobragem terminando o segmento de texto precedente como entrada e proporciona um movimento dos olhos mínimo aquando da leitura da cascata de frases. Outra forma de realização adiciona uma descida esquerda para pontos de dobragem de Classe 1 precedentes, e uma descida direita para pontos de dobragem de Classe 3 precedentes. Uma forma de realização preferida permite a especificação, pelo leitor, de deslocamento adicional direito ou esquerdo para os pontos de dobragem, incluindo valores introduzidos pelo leitor para pontos de dobragem primários, e cada classe e subclasse de pontos de dobragem secundários. Uma forma de realização armazena o deslocamento adicionado numa tabela em unidades de caracteres. Com o deslocamento horizontal determinado, permanece a apresentação do texto.

No exemplo da FIG. 9, a Super-frase 1, "*I had a farm in Africa*" está separada da Super-frase 2, "*at the foot of the Ngong Hills.*", pelo ponto de dobragem primário após a vírgula. A Super-frase 2 é dividida em duas Mini-frases pelo ponto de dobragem secundário antes da preposição "*of*". Na forma de realização ilustrada na FIG. 9, são apresentadas duas minifrases como o mesmo segmento de texto que a mini-frase "*in Africa*", o que é menos que o comprimento mínimo de três palavras especificado pelo leitor. A primeira Mini-frase, "*at the foot*",

numa nova linha é justificada ao centro abaixo do segmento de texto acima e desviada para a direita por um ângulo de descida direito. A última Mini-frase de "*of the Ngong Hills*", é desviada para a esquerda devido a uma regra de deslocamento horizontal exigindo o desvio para a esquerda quando a linha precedente é terminada com um ponto de dobragem secundário em vez de um primário. Assim, no exemplo da FIG. 9, o deslocamento horizontal total é determinado por uma combinação do ângulo de descida e da justificação da linha.

As marcas de pontuação emparelhadas incluindo parêntesis, parêntesis rectos, chavetas, e marcas de citação, servem como marcas de pontuação que determinam as posições de pontos de dobragem primários numa forma de realização preferida. Numa forma de realização, esta pontuação emparelhada é incluída como uma entrada para as regras de deslocamento horizontal, incluindo as regras de justificação horizontal e de ângulo de descida. Por exemplo, um segmento longo de texto com multi-frases contido entre parêntesis pode ter desvios horizontais reduzidos e deslocamento vertical reduzido, *i.e.*, menos espaçamento entre linhas do que o outro texto na área activa de apresentação da superfície do monitor.

No passo 112, é criado o texto melhorado. Neste passo, são criados os códigos necessários para criar um segmento de texto apresentado de forma correcta. Por exemplo, quando as especificações do leitor requerem a coloração de palavras técnicas em vermelho, e o texto enriquecido indica que uma palavra é uma palavra técnica, pode ser criada uma sequência de fuga que seja interpretada pelo passo de apresentação como requerendo texto em vermelho. Pode ser necessário um código

semelhante para a animação. O texto melhorado pode ser armazenado nesta altura para apresentação posterior.

No passo 113, o texto melhorado é apresentado no dispositivo monitor, um segmento de texto por cada linha formada de novo. O texto melhorado também pode incluir a animação, a cor do fundo, a cor do texto, a etiquetagem, e os ritmos de apresentação discutidos acima. Numa forma de realização preferida, a cor do fundo é apresentada em função das posições da frase e do parágrafo.

Noutra forma de realização, ilustrada na FIG. 11, algum texto é inicialmente apresentado numa cor ou brilho durante um período de tempo, seguido por uma apresentação numa segunda cor ou brilho. Na FIG. 11, a frase é apresentada em três linhas formadas de novo indicadas pelas setas 200, 202, e 203. Dentro destas linhas, o texto é apresentado todo na mesma linha, com as palavras da FIG. 11 mostradas em linhas diferentes para ilustrar a mudança temporal da cor/realce 206 do texto inicial para a cor/realce 207 de texto, e da cor/realce 208 do texto inicial para 209. O texto "Af" em 206 é a sílaba enfatizada em "Africa", e é inicialmente mostrado na cor/realce 206 por esta razão. O texto "gong" é também a sílaba enfatizada em "Ngong", e tem a cor/realce 208 inicial, seguida pela cor/realce 209.

A forma de realização preferida descrita na FIG. 1 permite a edição do texto enriquecido do passo 132 e do texto melhorado do passo 113. Numa forma de realização preferida, é utilizado um dispositivo apontador de palavra tal como um rato de computador para seleccionar uma porção de texto a ser editado. A frase enriquecida correspondendo a essa porção de texto melhorado é seleccionada internamente, e os atributos tornados disponíveis

para edição. Isto pode ser implementado utilizando menus permitindo que cada atributo do texto enriquecido e do texto melhorado seja editado. Nos casos em que as edições do texto enriquecido possam ter efeito sobre a apresentação, o texto é pré-processado novamente, de acordo com o passo 110. Na sessão de edição do passo 114, as anotações do leitor, sejam escondidas ou imediatamente visíveis, são aceites e podem ser armazenadas nas frases enriquecidas e melhoradas. Estas anotações desempenham a mesma função que as notas manuscritas num livro. As características editáveis incluem as partes de fala, a definição, a cor, o texto, os pontos de dobragem, e o deslocamento horizontal.

Os atributos editados e a apresentação desta forma de realização também podem ser armazenados, no passo 116, mantendo as mudanças. Numa forma de realização, apenas as frases melhoradas são armazenadas. Numa forma de realização preferida, tanto o texto enriquecido como o melhorado são armazenados. Em formas de realização preferidas, algumas edições são gravadas seja como específicas do leitor seja como globais, para apresentação a todos os leitores do texto. A edição manual do texto melhorado é especialmente útil no caso em que o mesmo texto seja visualizado repetidamente por outros, tal como num livro electrónico.

Numa variação da forma de realização da FIG. 1, os passos 113, 114, e 116 são omitidos, sem nenhuma intervenção humana e nenhuma apresentação imediata. Nesta forma de realização, é criado texto enriquecido e melhorado e armazenado para apresentação futura. O texto melhorado pode ser armazenado num formato de processamento de texto padrão, tal como os arquivos binários *Microsoft Word* (MARCA REGISTADA) ou *Corel Word Perfect*

(MARCA REGISTADA). Nesta forma de realização, o *software* de apresentação pode ser simples, pequeno, e rápido relativamente ao *software* necessário para consultar palavras e analisar o texto. Esta apresentação é uma forma de realização preferida para a distribuição em massa de texto melhorado para leitura como "livros electrónicos".

Numa variação relacionada da forma de realização da FIG. 1, o próprio texto melhorado é recuperado no passo 106, sem pré-processamento exigido e sem permitir edições. Esta apresentação é também uma forma de realização preferida para a distribuição em massa de texto melhorado para leitura como "livros electrónicos".

No que se refere à FIG. 1, passo 113, a apresentação de cada écran de texto melhorado pode ser activada manualmente por um toque no rato de computador. Numa forma de realização preferida, o ritmo de apresentação é controlado pelos parâmetros especificados pelo leitor incluindo o tempo de apresentação para o texto, e os tempos de chegada entre frases e entre parágrafos. Numa forma de realização muito preferida, é medido o Conteúdo do Texto, e o intervalo de apresentação do texto está dependente deste Conteúdo do Texto. Numa forma de realização preferida, o tempo de pronúncia do texto é utilizado como uma medida do Conteúdo do Texto para determinar o intervalo de apresentação. Noutra forma de realização, os pesos das frases são utilizados para medir o Conteúdo do Texto e para determinar o intervalo de apresentação.

Uma forma de realização preferida permite aos leitores aumentar o tempo de apresentação para uma frase, bem como acelerar e desacelerar o ritmo de apresentação. Este ritmo pode

ser gravado, tal como o comprimento de frase e a dificuldade da frase correspondendo a um intervalo de apresentação particular. A FIG. 10 ilustra um exemplo de um gráfico 3-D, um "Mapa da História de Amostragem do Intervalo de Apresentação", representando o tempo necessário para ler uma frase em função de dois atributos, a dificuldade da frase e o comprimento da frase. O tempo necessário para a leitura pode ser medido pelas preferências dos leitores ou pelo seguimento do movimento dos olhos mencionado previamente ou da actividade das ondas cerebrais. O tempo pode ser correlacionado com dois atributos, tal como na FIG. 10, ou qualquer número de atributos. O intervalo de apresentação numa forma de realização preferida é continuamente ajustado para corresponder ao tempo de leitura necessário previsto.

Em ainda outra forma de realização, mais apropriada quando o ritmo de leitura é medido rigorosamente, as entradas para as regras de dobragem são variadas e o ritmo de leitura resultante é seguido. Numa forma de realização adequada para criar um produto de leitura óptimo para o mercado de massas, as entradas para as regras de dobragem são variadas e os ritmos de leitura gravados para uma população amostra. As entradas são variadas para uma velocidade de leitura e compreensão óptimas. Por exemplo, a importância atribuída às regras de dobragem secundárias para partes de fala em relação ao peso da frase é variada para uma velocidade de leitura óptima.

Noutra forma de realização, a edição manual, pelo leitor, da apresentação inicial da frase é analisada quanto à contribuição relativa da estrutura da frase, partes de fala, comprimento, e complexidade. Após este "afinar" inicial ou período de "treino", são utilizados pesos similares nas regras

de dobragem para melhoramento automático do texto. Podem ser utilizadas edições subsequentes para refinar ainda mais as regras de dobragem.

No passo 118, a apresentação do texto pára quando não existe mais texto ou um leitor pede para parar. O passo 120 proporciona a limpeza incluindo o armazenamento do texto enriquecido e melhorado final, bem como o armazenamento do histórico da sessão de leitura.

Na descrição precedente foram expostas várias características e vantagens da invenção abrangidas por este documento. Deve compreender-se, contudo, que esta divulgação é, em muitos aspectos, apenas ilustrativa. Podem ser feitas mudanças nos detalhes, particularmente no que respeita a combinar, separar, e organizar os passos, sem exceder o âmbito desta invenção. O âmbito da invenção é, evidentemente, definido na linguagem em que são expressas as reivindicações anexadas.

Lisboa, 6 de Julho de 2007

REIVINDICAÇÕES

1. Método para melhorar a apresentação de um texto a partir de um texto em linguagem natural legível numa máquina, baseado em parâmetros específicos do leitor incluindo pelo menos as dimensões do campo de visualização compreendendo:

a análise do referido texto em pontuação e partes de fala para extracção de atributos específicos do texto;

o armazenamento dos referidos atributos específicos relativamente às partes de fala para produzir um texto enriquecido;

a aplicação de regras de dobragem primárias seguidas de regras de dobragem secundárias ao referido texto enriquecido, aplicadas por ordem segundo uma ordenação de regras de dobragem dividindo deste modo o referido texto em segmentos de texto tendo as referidas regras de dobragem, pelo menos, os referidos atributos de pontuação e os atributos de partes de fala como entradas e atributos visuais como saídas;

a aplicação de regras de dobragem secundárias até ser atingido um limite, sendo este limite, de um modo preferido, o comprimento mínimo de linha; e

em que os atributos visuais incluem a apresentação dos segmentos de texto em novas linhas;

a aplicação de regras de deslocamento horizontal dos segmentos de texto aos referidos segmentos de texto para determinar um deslocamento horizontal para cada segmento de texto, incluindo as referidas regras de deslocamento horizontal partes de fala como entradas e atributos visuais de deslocamento horizontal como saídas para produzir um texto melhorado; e

a apresentação do referido texto melhorado colocando em cascata os segmentos de texto em linhas para baixo e através dum monitor.

2. Método da reivindicação 1 em que os referidos atributos específicos do texto incluem a posição do texto dentro de um corpo de texto, a apresentação do referido texto inclui cor de fundo e o referido processamento de texto inclui a variação da referida cor de fundo de acordo com a referida posição do texto.
3. Método da reivindicação 1 em que os referidos atributos específicos do texto incluem uma medida da dificuldade do texto, e a referida apresentação do texto inclui um ritmo de avanço automático do texto, e o referido processamento de texto inclui a variação do referido ritmo de apresentação do texto de acordo com a referida medida de dificuldade do texto.
4. Método da reivindicação 3 em que a referida medida de dificuldade do texto inclui um tempo de pronúncia estimado do referido texto e/ou um nível educacional estimado do referido texto e/ou o número de sílabas no referido texto.

5. Método da reivindicação 1 em que os referidos atributos visuais incluem a apresentação dos referidos segmentos de texto numa cor dependendo das referidas partes de fala e/ou a apresentação dos referidos segmentos de texto em novas linhas.
6. Método da reivindicação 1, em que

as referidas regras primárias incluem a pontuação como entradas;

as referidas regras secundárias incluem partes de fala como entradas.
7. Método da reivindicação 6 em que as referidas regras de dobragem incluem ainda uma microgramática para que as referidas partes de fala não sejam ambíguas.
8. Método da reivindicação 1 compreendendo ainda a determinação das partes de fala prováveis das referidas palavras no referido passo de análise.
9. Método da reivindicação 9 compreendendo ainda:

a provisão de regras de apresentação de cores, em que as referidas regras de apresentação de cores incluem as referidas partes de fala prováveis como entradas e a cor do texto como uma saída; e

a apresentação do referido texto de acordo com a referida cor de saída do texto.

10. Método da reivindicação 1 e 6, em que a aplicação de regras de dobragem primárias e secundárias compreende:

proporcionar um comprimento mínimo de segmento de texto;

proporcionar um comprimento máximo de segmento de texto;

determinar as posições dos pontos de dobragem primários no referido texto aplicando as regras de dobragem primárias, dividindo desse modo o referido texto em super-frases;

determinar as posições dos pontos de dobragem secundários nas super-frases aplicando as regras de dobragem secundárias, dividindo desse modo as super-frases em segmentos de texto;

repetir o passo anterior até que todos os segmentos de texto não sejam maiores do que o comprimento máximo do segmento de texto, e não sejam menores do que o comprimento mínimo do texto.

11. Método da reivindicação 1 ou 10 ccompreendendo ainda

aceitar as edições do utilizador dos pontos de dobragem de texto apresentado; e

armazenar as edições do utilizador numa forma legível numa máquina; e/ou

aceitar as edições do utilizador das partes de fala do texto apresentado; e armazenar as edições do utilizador numa forma legível numa máquina; e/ou

aceitar as anotações do utilizador às partes de fala do texto apresentado; e armazenar as anotações do utilizador numa forma legível numa máquina.

12. Dispositivo para melhorar a apresentação de um texto a partir de um texto em linguagem natural legível numa máquina, baseado em parâmetros específicos do leitor incluindo pelo menos, as dimensões do campo de visualização do texto compreendendo:

meios de análise para analisar o referido texto em pontuação e partes de fala para extracção de atributos específicos do texto;

meios de armazenamento para armazenar os referidos atributos específicos do texto relativamente às partes de fala para produzir um texto enriquecido;

meios para aplicação de regras de dobragem primárias seguidas de regras de dobragem secundárias ao referido texto enriquecido, aplicadas por ordem segundo uma ordenação de regras de dobragem dividindo desse modo o referido texto em segmentos de texto, tendo as referidas regras de dobragem, pelo menos, os referidos atributos de pontuação e os atributos de partes de fala como entradas e atributos visuais como saídas;

meios para aplicação de regras de dobragem secundárias até ser atingido um limite, sendo este limite, de um modo preferido, de um modo preferido, o comprimento mínimo de linha; e

em que os atributos visuais incluem a apresentação dos segmentos de texto em novas linhas;

meios para aplicação de regras de deslocamento horizontal dos segmentos de texto aos referidos segmentos de texto para determinar um deslocamento horizontal para cada segmento de texto, incluindo as referidas regras de deslocamento horizontal partes de fala como entradas e atributos visuais de deslocamento horizontal como saídas para produzir um texto melhorado; e

meios para apresentação do referido texto melhorado colocando em cascata os segmentos de texto em linhas para baixo e através de um monitor.

13. Dispositivo da reivindicação 12 em que os referidos meios para aplicação de regras de dobragem primárias e secundárias compreendem:

meios para a introdução dum comprimento mínimo de segmento de texto, para introduzir um comprimento mínimo de segmento de texto;

meios para a introdução dum comprimento máximo de segmento de texto, para introduzir um comprimento máximo de segmento de texto;

meios de determinação primária para determinar as posições dos pontos de dobragem primários no referido texto aplicando as regras de dobragem primárias, dividindo desse modo o referido texto em super-frases;

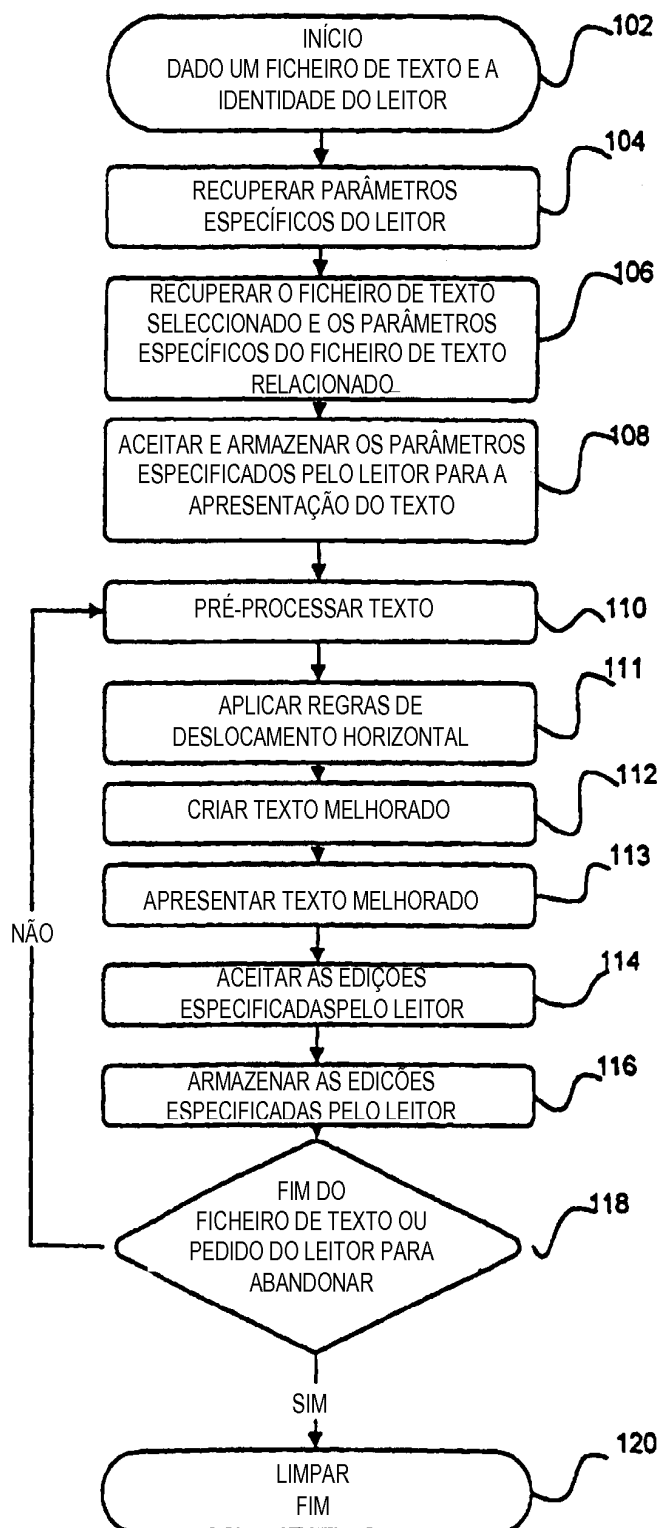
meios de determinação secundária para determinar as posições dos pontos de dobragem secundários nas super-frases aplicando as regras de dobragem secundárias, dividindo desse modo as super-frases em segmentos de texto;

meios para repetir o referido passo para determinar as posições dos pontos de dobragem secundária até que todos os segmentos de texto não sejam maiores do que o comprimento máximo do segmento de texto, e não sejam menores do que o comprimento mínimo do texto.

14. Dispositivo da reivindicação 13 em que o texto enriquecido inclui palavras e frases e é implementado como uma lista ligada de nós, possuindo cada nó atributos de palavra e de frase.
15. Dispositivo da reivindicação 12, em que as referidas regras de dobragem incluem regras primárias e regras secundárias, as referidas regras primárias incluem a pontuação como entradas e as referidas regras secundárias incluem partes de fala como entradas; compreendendo uma tabela contendo conjuntos de palavras estabelecendo uma hierarquia de pontos de dobragem secundários.

Lisboa, 6 de Julho de 2007

Fig. 1



N° SE	Elemento	Tipo	Cat. PM	Pos. da palavra no Dic.	Pos. da palavra na MG	Maíscula	N° caracteres	N° de sílabas	Tempo de pronúncia, segundos	Nível educacional
1	I	w		Pron.		sim	1	1	,5	1
2		s								
3	had	w		verbo			3	1	,25	1
4		s								
5	a	w		artigo			1	1	,25	1
6		s								
7	farm	w		n-v-a	Subst.		4	1	,75	1
8		s								
9	in	w		Prep.			2	1	,25	1
10		s								
11	Africa	w		Subst.		sim	6	3	1,0	3
12	,	pm	virgula				1		,5	
13		s							,5	
14	at	w		Prep.			2	1	,25	1
15		s								
16	the	w		artigo			3	1	,25	1
17		s								
18	foot	w		n-v-a	Subst.		4	1	,5	1
19		s								
20	of	w		Prep.			2	1	,25	1
21		s								
22	the	w		artigo			3	1	,25	1
23		s								
24	Ngong	w		Subst.		sim	5	2	1,0	12
25		s								
26	Hills	w		Subst.		sim	5	1	,75	2
27	.	pm	Ponto final	150	152				,5	

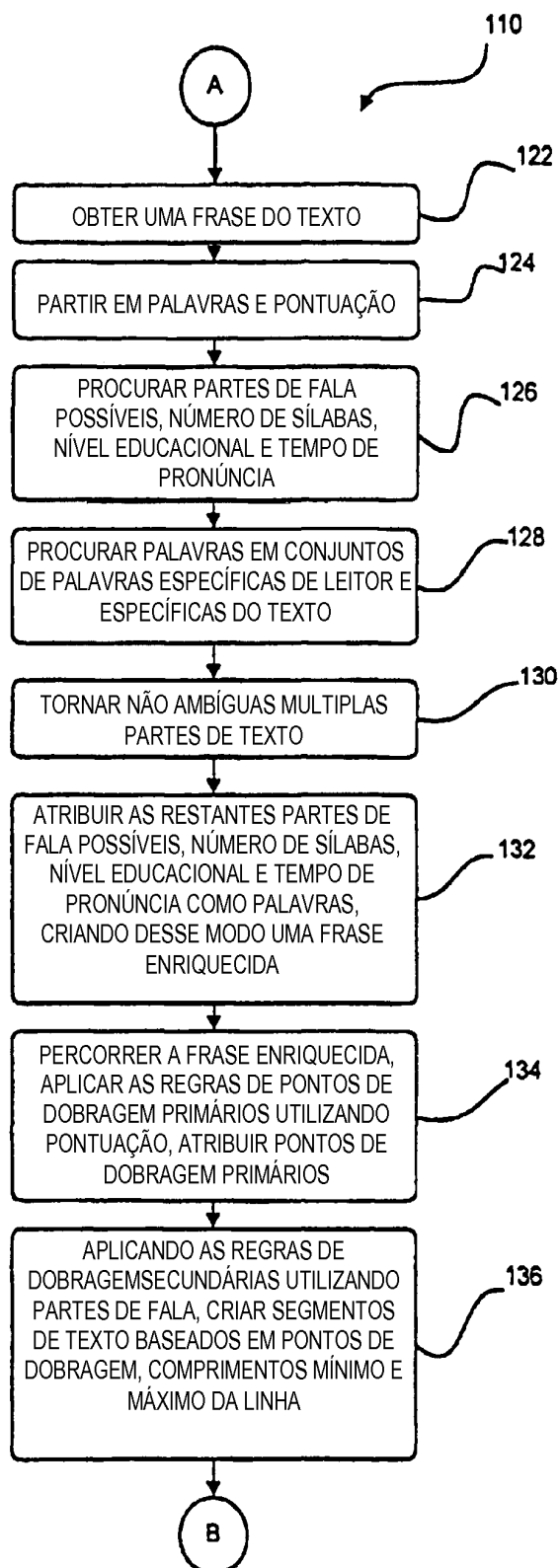
Fig. 2

The geographical position,
and the height of the land,
combined to create a landscape
that had not its like
in all the world.

Fig. 3

The geographical position,
and the height
of the land,
combined to create a landscape
that had not its like
in all the world.

Fig. 4



CLASSE 1,SUBCLASSE 1 PREPOSIÇÕES	AMID, BESIDES, AGAINST, AMONG, ABOUT, BY (COM GERÚNDIO), DURING, SINCE, THROUGHOUT, WITHIN, UNTO, ALON, EXCPT, AROUND, THROUGH, UNTIL, WITOUT
CLASSE 1,SUBCLASSE 2 CONJUNÇÕES	ALTHOUGH, AS THOUGH, UNLESS, UNTIL, AS IF, SO THAT, IN ORDER THAT, BECAUSE, BUT
CLASSE 1,SUBCLASSE 3 PRONOMES RELATIVOS	TO WHOM, TO WHOEVER
CLASSE 1,SUBCLASSE 4 OUTROS	THATER THAN, IN LIEU OF, INSTEAD OF
CLASSE 2,SUBCLASSE 1 PREPOSIÇÕES	ABOVE, BEFORE, BENEATH, BETWEN, UP, ACROSS, TOWARD, UPON, AFTER, UNDER, TO (COM VERBO)
CLASSE 2,SUBCLASSE 2 CONJUNÇÕES	WHEN, WHERE, WHENEVER, WHEREVER, WHILE, AND, (COM RESTRIÇÕES), IF
CLASSE 2,SUBCLASSE 3 PRONOMES RELATIVOS	WHAT, WHOM, WHICH, WHO, WHOSE
CLASSE 2,SUBCLASSE 4 OUTROS	INTERFACE SUBSTÂNTIVO/VERBO, ADVÉBIOS A SEGUIR A UM VERBO
CLASSE 3,SUBCLASSE 1 PREPOSIÇÕES	FOR, AT, TO (COM NÃO-VERBOS), WITH, OF, IN, BY (SEM GERÚNDIO), FROM, ON

Fig. 6

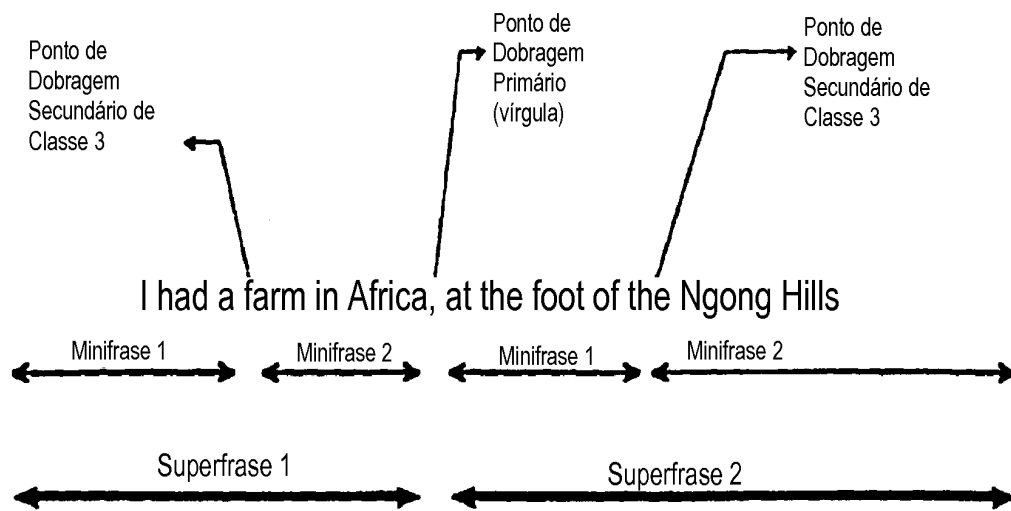


Fig. 7

Fig. 8

```
//SP = SUPERFRASE
//SFR = REGRA DE DOBRAGEM SECUNDÁRIA
//MIN = COMPRIMENTO MÍNIMO DE LINHA
//MAX = COMPRIMENTO MÁXIMO DE LINHA

PARA TODA AS SP NA FRASE
  PARA TODOS OS NÍVEIS DE SFR, DE PEQUENO A GRANDE
    PARA TODAS AS PALAVRAS NA SP, DA ESQUERDA PARA A DIREITA
      SE O SEGMENTO DE TEXTO FOR > MIN
        APLICAR SFR, CRIANDO UM NOVO SEGMENTO DE TEXTO
      CASO CONTRÁRIO
        NÃO APLICAR SFR ASSIM\
        ADICIONANDO AO COMPRIMENTO DO PRÓXIMO\
        SEGMENTO DE TEXTO
      TERMINAR CONDIÇÃO
    TERMINAR PARA TODAS AS PALAVRAS NA SF
  TERMINAR PARA TODOS OS NÍVEIS DE SFR
TERMINAR PARA TODOS OS SP NA FRASE

PARA TODOS OS SEGMENTOS DE TEXTO CRIADOS
  ENQUANTO O COMPRIMENTO DO SEGMENTO DE TEXTO FOR > MAX
    APLICAR A REGRA DE COLAPSO ASSIM\
    CRIANDO MÚLTIPLOS SEGMENTOS DE TEXTO
  TERMINAR ENQUANTO
TERMINAR PARA TODOS SEGMENTOS DE TEXTO CRIADOS
```

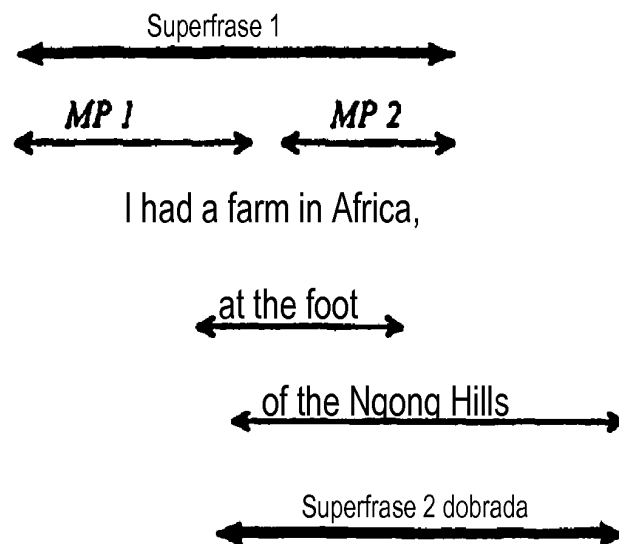


Fig. 9

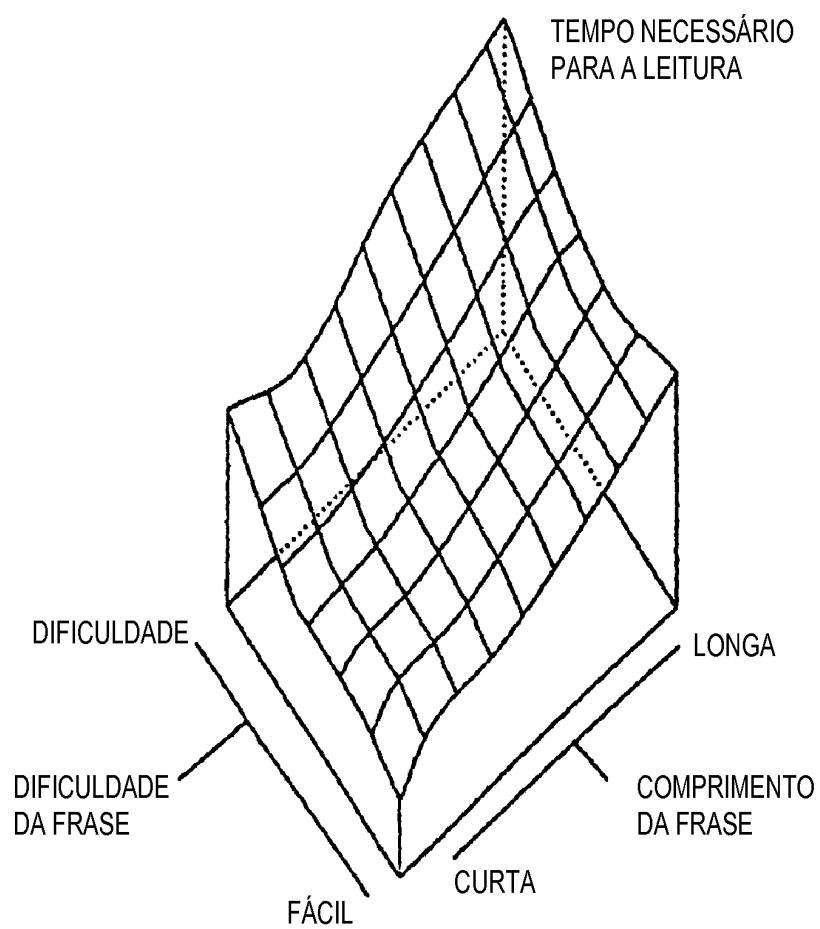


Fig. 10

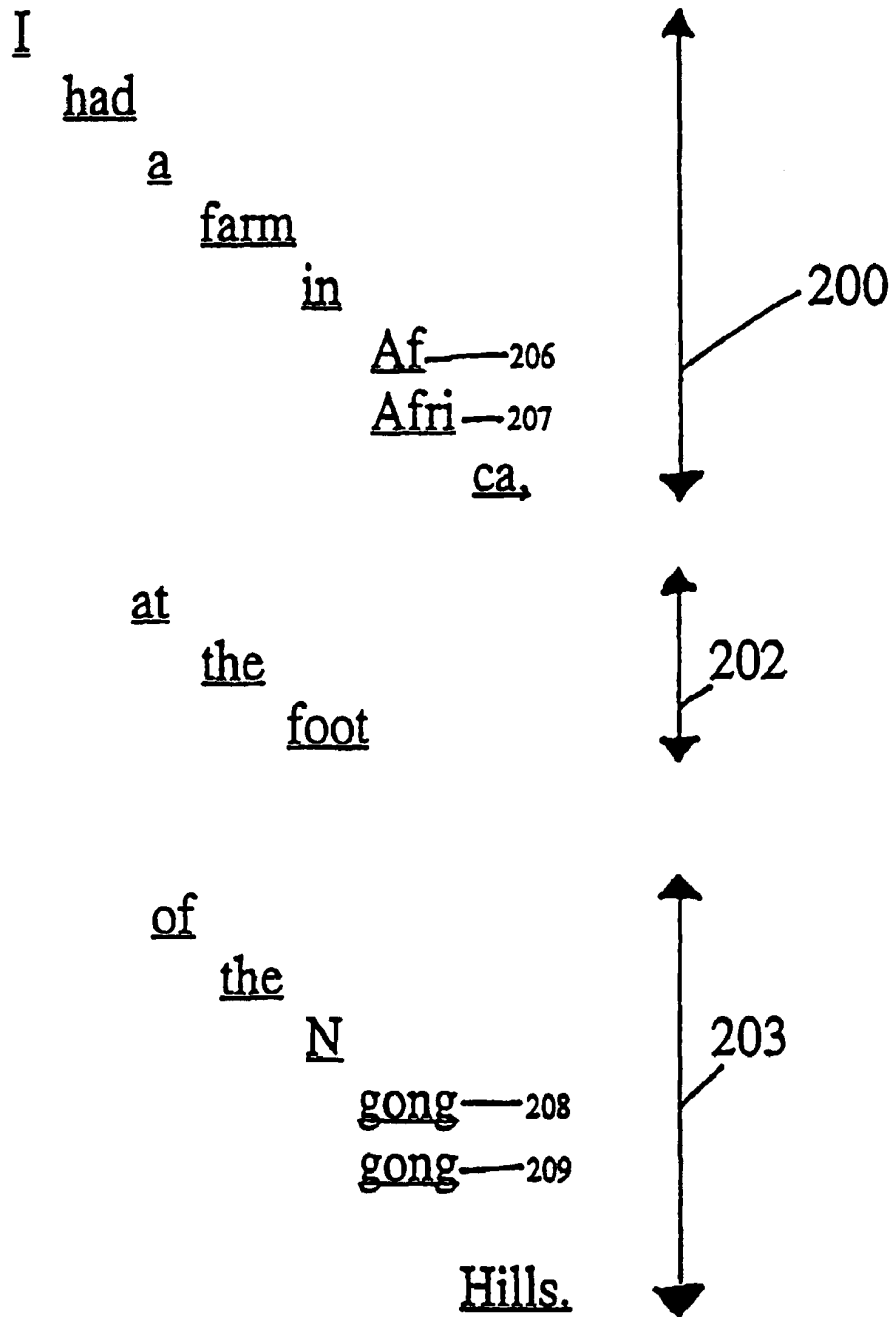


Fig. 11