

**INSTITUTO DE ENSINO SUPERIOR DE RONDÔNIA  
BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

**FONOMOUSE: FACILITADOR DO USO DO COMPUTADOR PARA  
PORTADORES DE NECESSIDADES ESPECIAIS**

**CLEYTON CESAR FERRARI**

**ARIQUEMES, 2004**

**FONOMOUSE: FACILITADOR DO USO DO COMPUTADOR PARA  
PORTADORES DE NECESSIDADES ESPECIAIS**

Trabalho de Graduação apresentado ao Curso de Sistemas de Informação, do Instituto de Ensino Superior de Rondônia, intitulado FonoMouse: Facilitador do Uso do Computador para Portadores de Necessidades Especiais, pelo aluno Cleyton Cesar Ferrari, sob orientação do professor Marcelo Douglas.

Ariquemes, Novembro de 2004.

**FONOMOUSE: FACILITADOR DO USO DO COMPUTADOR PARA  
PORTADORES DE NECESSIDADES ESPECIAIS**

Cleyton Cesar Ferrari

Trabalho aprovado como requisito parcial para obtenção do grau de Bacharel em Sistemas de Informação no Curso de Graduação em Sistemas de Informação do Instituto de Ensino Superior de Rondônia, pela comissão formada pelos professores:

Comissão examinadora:

---

Prof. Orientador Marcelo Douglas  
FAAr – Faculdades Associadas de Ariquemes

---

Prof. \_\_\_\_\_  
FAAr – Faculdades Associadas de Ariquemes

---

Prof. \_\_\_\_\_  
FAAr – Faculdades Associadas de Ariquemes

Ariquemes, 22 de Novembro de 2004.

## AGRADECIMENTOS

Agradeço,

Especialmente ao meu amigo e companheiro de turma Anderson, que na hora mais difícil desta caminhada não me deixou desanimar, sempre com pensamentos positivos e com uma frase certa “Nós vamos terminar!” e por fim, Terminamos!

A minha esposa que sempre esteve presente e soube esperar pacientemente que eu terminasse estas noites de estudo longe dela.

A minha mãe que sempre me apoiou, e me incentivou e que um dia na formatura de meu tio disse “que o meu sonho é ver um filho formado!”, pois é mãe agora a senhora pode comemorar!

E a todos que direta ou indiretamente contribuíram para que eu pudesse realizar esta jornada, o meu muito obrigado.

## DEDICATÓRIA

*“Dedico este trabalho a todos que me acompanharam nesta caminhada, em especial aos meus colegas de turma que foram ficando pelo caminho durante esta jornada, e que não desistam de lutar nunca, pois o que você realmente quer você consegue!”.*

## RESUMO

A execução e aplicação do projeto têm por finalidade o desenvolvimento de uma ferramenta que possibilite aos portadores de necessidades especiais movimentarem o ponteiro do mouse na tela utilizando-se da voz como único método de interação com o computador. Demonstra a necessidade do uso do computador para os portadores de necessidades especiais, informando de maneira fácil o que vem sendo aplicado e feito para inserção dos portadores de necessidades especiais no mundo da informação digital. Apresenta os métodos utilizados no desenvolvimento do sistema de movimentação do mouse, bem como sua modelagem de dados, utilizando-se de todos os diagramas determinados pela UML. São apresentadas de forma clara e objetiva, as funcionalidades que o sistema possui, utilizando-se de quadros, figuras e telas prontas do sistema. São apresentados logo no início do trabalho os conceitos da linguagem natural, onde é explanado sobre suas facilidade de uso e sua ligação ao homem, além dos demais conceitos sobre portadores de necessidade especiais, e a interface natural de reconhecimento de voz. Em suma, mostra a utilização da base de dados de reconhecimento de voz disponível no mercado, a sua integração com a ferramenta de desenvolvimento de software que mais se adapta a máquina de reconhecimento de voz, sendo no final do trabalho apresentado o sistema proposto.

Palavras-chave: Reconhecimento de voz; Portador de necessidades especiais; Interface Natural; Mouse.

## ABSTRACT

The execution and application of the project have for purpose the development of a tool that makes possible the special carriers of necessities to put into motion the hand of mouse in the screen using itself of the voice as only method of interaction with the computer. It demonstrates the necessity of the use of the computer for the special carriers of necessities, informing in easy way what it comes being applied and done for insertion of the special carriers of necessities in the world of the digital information. It presents the methods used in the development of the system of movement of mouse, as well as its modeling of data, using it of all the diagrams determined for the UML. Are presented of clear and objective form, the functionalities that the system possesses, using itself of pictures, figures and ready screens of the system. The concepts of the natural language are presented soon in the beginning of the work, where it is speaking on its easiness of use and it's linking to the man, beyond the too much concepts on special carriers of necessity, and the natural interface of voice recognition. In short, it shows the use of the database of recognition of available voice in the market, its integration with the tool of software development that more if to adapt schemes it of voice recognition, being in the end of the presented work the considered system.

Word-key: Recognition of voice; Carrier of special necessities; Natural Interface; Mouse.

## SUMÁRIO

INTRODUÇÃO.....	13
1. PORTADOR DE NECESSIDADES ESPECIAIS .....	16
1.1. Categorias de Deficiências .....	16
1.2. Exclusão digital .....	18
1.3. Necessidade do Uso do Computador.....	19
2. LINGUAGEM NATURAL.....	21
2.1. Interface homem-computador.....	21
2.2. Interface em Linguagem natural.....	23
3. RECONHECIMENTO DE VOZ .....	26
3.1. Reconhecimento de Voz.....	27
3.2. Funcionamento do Mecanismo de Reconhecimento de Voz .....	29
4. CONCEITOS GERAIS DA METODOLOGIA UTILIZADA .....	33
4.1. Método Utilizado.....	33
4.2. Análise Orientada a Objetos .....	33
4.2.1. Objetos.....	33
4.2.2. Encapsulamento.....	34
4.2.3. Classes e Instâncias .....	34
4.2.4. Herança.....	35
4.2.5. Polimorfismo .....	35
4.3. UML ( <i>Unified Modeling Language</i> ) .....	36
4.3.1. Diagrama de Caso de Uso .....	36
4.3.2. Diagrama de Classes.....	36
4.3.3. Diagrama de Interação (seqüência ou colaboração).....	37
4.3.4. Diagrama de Estado e de Atividade .....	37
4.3.5. Diagrama de Componentes e de implantação.....	38
4.4. Ambiente de Desenvolvimento Delphi .....	38
4.5. Microsoft Speech.....	40
4.6. Modelagem de Dados do FonoMouse .....	40
4.6.1. Diagrama de Caso de Uso do FonoMouse .....	40
4.6.2. Diagrama de Classes do FonoMouse .....	41
4.6.3. Diagrama de Interação do FonoMouse.....	41
4.6.4. Diagrama de Estado e de Atividade do FonoMouse .....	42
4.6.5. Diagrama de Componente e de implantação do FonoMouse.....	43
5. FUNCIONALIDADES DO SISTEMA .....	44



5.1.	Tela de carregamento.....	44
5.2.	Tela Principal.....	44
5.3.	Painel de Configurações .....	45
5.3.1.	Configurar Mouse.....	45
5.3.2.	Configurar Grade.....	46
5.3.3.	Configurar voz e microfone.....	47
5.4.	Ajuda .....	48
5.5.	Sobre o FonoMouse.....	49
5.6.	Requisitos de instalação do FonoMouse .....	49
CONCLUSÃO.....		51
REFERÊNCIAS BIBLIOGRÁFICAS .....		53
MARCAS REGISTRADAS .....		56
ANEXOS.....		57
	Projeto de Monografia.....	58

## LISTA DE TABELAS

Tabela 1- Número de Portadores de deficiência: Por tipo de deficiência. ....	18
--	----

## LISTA DE QUADROS

Quadro 1 - Categorias de Deficiências e suas descrições.....	17
Quadro 2 - Formas que se apresenta a deficiência física.....	17

## LISTA DE FIGURAS

Figura 1 - Diagrama de Casa de Uso .....	36
Figura 2 - Diagrama de Classes .....	36
Figura 3 - Diagrama de Interação .....	37
Figura 4 - Diagrama de Atividade .....	37
Figura 5 - Diagrama de Componentes e de Implantação.....	38
Figura 6 - Ambiente de Desenvolvimento do Delphi.....	39
Figura 7 - Paleta de Componentes Speech SDK instalado no Delphi .....	40
Figura 8 - Diagrama de Caso de Uso do FonoMouse .....	40
Figura 9 - Diagrama de Classe do FonoMouse .....	41
Figura 10 - Diagrama de Interação do FonoMouse .....	42
Figura 11 - Diagrama de Estado e de Atividade do FonoMouse.....	43
Figura 12 - Diagrama de Componente e de Implantação do FonoMouse .....	43
Figura 13 - Tela de Carregamento do FonoMouse.....	44
Figura 14 - Tela Principal do Sistema .....	44
Figura 15 - Painel de Configurações .....	45
Figura 16 - Configurações do Mouse .....	45
Figura 17 - Exibição da Grade.....	46
Figura 18 - Configurações de Grade .....	46
Figura 19 - Configurações de Voz.....	47
Figura 20 - Configuração de Voz e Microfone.....	47
Figura 21 - Assistente de Configuração do Microfone.....	48
Figura 22 - Sistema de Ajuda do FonoMouse .....	48
Figura 23 - Sobre o FonoMouse .....	49

## SIGLAS E ABREVIATURAS

DLL – *Dinamic Link Library* (Biblioteca de Ligação Dinâmica).

IBGE – Instituto Brasileiro de Geografia e Estatística.

IBM – *Industrial Business Machines* (Indústria e Negócios de Máquinas).

IDE – *Integrated Development Environment* (Ambiente de Desenvolvimento Integrado).

MB – Megabytes.

OLE – *Object Linking and Embedding* (Encaixe e Ligamento de Objetos).

OMS – Organização Mundial de Saúde.

PUC – Pontifícia Universidade Católica.

RAD – *Rapid Application Development* (Desenvolvimento Rápido de Aplicação).

SDK – *Software Develop Kit* (Conjunto de Ferramenta de Desenvolvimento de Software).

TCC – Trabalho de Conclusão de Curso.

UML – *Unified Modeling Language* (Linguagem de Modelagem Unificada).

USB – *Universal Serial Bus*.

## INTRODUÇÃO

Existem muitas formas de comunicação entre os programas e os usuários de computadores, por exemplo: menus, linguagem de comandos, voz, linguagem natural, entre outras.

Dentre estas, a interface gráfica que se utiliza de periféricos de entrada como mouse e teclado para a interação do usuário com o sistema ou ainda, por meio de teclas de função, como exemplo: F2 – salva, CLS - limpa tela, etc., trazem consigo a rigidez de sintaxe, deixando excluído do uso do sistema portador de necessidades especiais que não possuam a movimentação dos membros superiores.

Sendo assim, o que este projeto se propõe é um meio de se fazer a inclusão digital destes portadores de necessidades especiais, que até então de certa forma se encontram excluídos digitalmente.

A tecnologia de reconhecimento de voz tornou-se disponível em meados dos anos 90, mas só agora se está consolidando comercialmente, graças a uma excelente relação custo-benefício. Esse fato está fortemente atrelado a algumas características que tornaram o emprego da tecnologia altamente atrativa. (NIKOS, 2004).

O funcionamento básico do reconhecimento de voz resulta da integração entre uma placa com processadores digitais que realiza a captura e o tratamento do áudio falado, com um algoritmo especializado que quebra essa informação falada em pequenos pedaços batizados de fonemas. Assim, cada som individual pode ser identificado e comparado a uma lista pré-definida de palavras ou frases. (PAULA, 2000).

Entre as vantagens desse tipo de plataforma em relação a outras tecnologias existentes, a mais imediata é o fato da fala ser inerente ao ser humano. Logo, por paralelismo, os sistemas de reconhecimento de voz mostram-se como interfaces naturais mais amigáveis aos portadores de necessidades especiais. (HÜBNER, 2000).

A automação do atendimento pode ser empregada em uma série de aplicações. Na área de entretenimento e informação é possível constatar a criação de produtos que disponibilizam conteúdos, utilizando menus de fácil navegação e grande interatividade entre usuário e sistema. Como exemplos, existem a leitura de e-mails, acesso a agendas, discagem ativada por voz, horóscopo, previsão de tempo e notícias.

Outro tipo de aplicação é o acesso à informação, na qual o usuário pode saber, por exemplo, os horários de vôos e partidas de ônibus e aviões. Além disso, a tecnologia de reconhecimento de voz pode ser usada em produtos bancários, preenchimento de fichas de visitas para vendedores e outros serviços que registram e consultam informações em bases de dados.

Por fim, as possibilidades do reconhecimento de voz têm levado uma série de empresas, de diversos setores, a optar por essa tecnologia em seus sistemas de atendimento. Na prática, essa interface pode resultar em inúmeros desdobramentos positivos tais como redução de custos operacionais, criação de novos produtos e aumento de receitas para as empresas (HÜBNER, 2000).

Na abordagem da ferramenta a ser desenvolvida, não será demonstrada a criação da máquina de reconhecimento de voz, mas a utilização de uma já existente, ao passo que o foco principal é o desenvolvimento de uma ferramenta que possibilite a interação de usuários que de alguma forma estão impossibilitados de movimentar os membros superiores, com este mundo brilhante da computação.

Este trabalho está disposto em 5 Capítulos e foi adotado o seguinte roteiro de apresentação:

Capítulo 1 apresenta os conceitos sobre os portadores de necessidades especiais, bem como suas categorias, passando pela exclusão digital e a necessidade que os portadores de

necessidades especiais têm na utilização do computador para se integrarem a este mundo digital.

Capítulo 2 demonstra de uma maneira clara o conceito da linguagem natural e da interface homem-computador.

Capítulo 3 faz uma introdução sobre o reconhecimento de voz, explicando de maneira objetiva o funcionamento do mecanismo de reconhecimento de voz.

Capítulo 4 apresenta os conceitos gerais da metodologia utilizada, explicando-as de forma sucinta, além da modelagem e do ambiente de desenvolvimento utilizado na construção do sistema.

Capítulo 5 demonstra as funcionalidades do sistema, assim como os requisitos necessários para a instalação do sistema proposto.

E por fim a Conclusão, que visa demonstrar como o autor vem encerrar a elaboração deste trabalho.



## 1. PORTADOR DE NECESSIDADES ESPECIAIS

Pessoa portadora de deficiência é toda pessoa com perda ou anormalidade de uma estrutura ou função psicológica, fisiológica ou anatômica que acarrete incapacidade para o desempenho da atividade, dentro do padrão considerado normal para o ser humano. (BRASIL, 1999).

No entanto, de acordo com a PUC Minas (2000, p.18) “a terminologia atualmente aceita, para referirem-se as pessoas portadoras de algum tipo de deficiência, quer sensitiva, quer física ou mental, é pessoas portadoras de necessidades especiais”.

O Brasil possui uma enorme população de portadores de deficiência – 16 milhões de pessoas. No entanto, são pouquíssimos os que trabalham no mercado formal. Segundo estimativas disponíveis, dos 16 milhões de portadores de deficiência, 9 milhões estão em idade de trabalhar. Destes, apenas 2% trabalham no mercado formal, enquanto nos países mais avançados essa proporção fica entre 30% e 45% (MIRANDA, 2004, p. 1).

Entre os portadores de Necessidades Especiais, os que mais se destacam para a futura utilização do Sistema são os que de alguma forma estão impossibilitados de realizarem a movimentação ou utilização dos membros superiores.

### 1.1. Categorias de Deficiências

As categorias de classificação dos portadores de necessidades especiais existentes são: Física, Sensorial que se divide em auditiva e visual, Deficiência Mental e, Deficiência Múltipla.

Quadro 1 - Categorias de Deficiências e suas descrições

<b>Deficiência</b>	<b>Descrição</b>
Física	Alteração completa ou parcial de um ou mais segmentos do corpo humano, acarretando o comprometimento da função física, exceto as deformidades estéticas e as que não produzem dificuldades para o desempenho de funções.
Sensorial	<i>Auditiva</i> – Perda parcial ou total das possibilidades auditivas sonoras, variando de graus e níveis (de 25 a 40 decibéis – Surdez leve; de 41 a 55 decibéis – Surdez moderada; de 56 a 70 decibéis – Surdez acentuada; de 71 a 90 decibéis – Surdez severa; acima de 91 decibéis – Surdez profunda). <i>Visual</i> – Perda ou a redução de capacidade visual em ambos os olhos, em caráter definitivo e que não possa ser melhorada ou corrigida com o uso de lentes e tratamento clínico ou cirúrgico. Entre os deficientes visuais têm-se os portadores de cegueira e os de visão subnormal.
Mental	Funcionamento intelectual significativamente inferior à média, apresentando níveis de comprometimento leve, moderado, severo ou profundo no comportamento adaptativo, tanto maior quanto o grau de comprometimento (dificuldade cognitiva).
Múltiplas	Associação de duas ou mais deficiências.

Fonte: PUC Minas, 2000.

Dentre estas categorias, a deficiência física se apresenta da seguinte forma: Paraplegia, Paraparesia, Monoplegia, Monoparesia, Tetraplegia, Tetraparesia, Triplegia, Triparesia, Hemiplegia, Hemiparesia, Paralisia Cerebral, Membros com deformação congênita ou adquirida.

Quadro 2 - Formas que se apresenta a deficiência física

<b>Formas</b>	<b>Descrição</b>
Paraplegia	Perda total das funções motoras dos membros inferiores.
Paraparesia	Perda parcial das funções motoras dos membros inferiores.
Monoplegia	Perda total das funções motoras de um só membro (podendo ser membro superior ou inferior).
Monoparesia	Perda parcial das funções motoras de um só membro (podendo ser membro superior ou inferior).
Tetraplegia	Perda parcial das funções motoras dos membros inferiores e superiores.
Tetraparesia	Perda parcial das funções motoras dos membros inferiores e superiores.
Triplegia	Perda total das funções motoras em três membros.
Triparesia	Perda parcial das funções motoras em três membros.
Hemiplegia	Perda total das funções motoras de um hemisfério do corpo (direito ou esquerdo).
Hemiparesia	Perda parcial das funções motoras de um hemisfério do corpo (direito ou esquerdo).
Amputação ou ausência de membros	Perda total de um determinado segmento de um membro (superior ou inferior).
Paralisia cerebral	Lesão de uma ou mais áreas do sistema nervoso central, tendo como consequência alterações psicomotoras podendo, ou não, causar deficiência mental.
Membros com deformação congênita ou adquirida	Anomalia física desde o nascimento ou adquirida.

Fonte: PUC Minas, 2000.

A ausência ou perda do membro pode ocorrer por diversas causas, congênita ou adquirida. As adquiridas são mais frequentes sendo os traumatismos as principais causas de amputação do membro superior (BRITO, 2004).

Tabela 1- Número de Portadores de deficiência: Por tipo de deficiência.

<b>DEFICIÊNCIA</b>	<b>BRASIL</b>
Deficientes Mentais (5% da população)	8.000.000
Deficientes Físicos (2% da população)	3.200.000
Deficientes Auditivos (1,5% da população)	2.400.000
Deficientes Múltiplos (1,0% da população)	1.600.000
Deficientes Visuais (0,5% da população)	800.000
Total (10% da população)	16.000.000

Fonte: (Com base em dados populacionais do Censo Demográfico 2000 – IBGE, apud MIRANDA, 2004).

## 1.2. Exclusão digital

A exclusão digital é mais uma consequência das diferenças sociais, econômicas e políticas já existentes de distribuição de poder e renda. Pode ser entendido como a situação na qual um indivíduo ou grupo de pessoas se encontra impossibilitados de utilizar as mais recentes tecnologias digitais. Surge daí a divisão digital, onde pessoas passam a ficar divididas em dois grupos: o dos que participa do mundo digital e o dos que ficam a parte (BORTOLON, 2004).

Como mostra o Quadro 2, existe algumas formas de deficiência física que impossibilitam a movimentação parcial ou total dos membros superiores. Podendo esta movimentação ser leve, e com alguma ajuda ser possível comer e escrever, mas pode ser tão severa que impeça completamente a movimentação, sendo a pessoa obrigada a contar com o auxílio de outras pessoas para realizar sua alimentação, higiene, acesso a itens de cultura, acesso a itens de comunicação assim por diante. Ficando assim de certa forma impossibilitados de se utilizarem das tecnologias disponíveis como é o caso do computador, o que de certo ponto causa a exclusão digital (PROJETO, 2002).

De qualquer forma, estudar para um Portador de Necessidades Especiais é sempre muito difícil, pois ele dependerá da ajuda constante de outras pessoas para escrever e ler. Essas dependências acabam por trazer ao longo do tempo o afastamento do portador de necessidades especiais do universo social, ou melhor, dizendo do universo digital. Apesar

dessa perspectiva tão negativa, hoje já é possível o acesso a um imenso arsenal tecnológico que tem tornado viável uma vida muito menos penosa, tanto para a pessoa portadora de necessidades especiais quanto pela sua família e a sociedade na qual ele está inserido. Muitos destes sistemas são ativados por voz, eliminando a necessidade do uso das mãos. Em alguns casos já estão disponíveis alternativas pouco convencionais como o acionamento de dispositivos com um sopro ou até pelo reconhecimento eletrônico do movimento do olho (PROJETO, 2002).

### 1.3. Necessidade do Uso do Computador

Segundo Valente (1999, p. 3), “O computador pode ser um importante recurso para promover a passagem da informação ao usuário ou facilitar o processo de construção de conhecimento”.

O grande avanço tecnológico atual, as redes de computadores, em especial a Internet, que permite conectar pessoas espalhadas pelo mundo todo, têm sido o novo impulso e a nova promessa em direção ao uso da tecnologia de computadores e da consciência de sermos “cidadãos do mundo” (VALENTE, 1999, p. 5).

Em todas as grandes empresas ou mesmo na idade da pedra, o ser humano precisou: anotar, contar, dividir, comparar seus pertences. Os grandes impérios tinham suas áreas administrativas que faziam esse serviço. Fazia-se anotando em pedras, couro, papel... Até a mais fantástica obra prima da humanidade que foi o computador. Muito tempo se passou até chegar a ele, mas depois dele, os grandes desenvolvimentos estão se realizando em períodos muito mais curtos! A utilização do computador ligado a outro e esse outro em mais outro, isso é, em rede, fez com que se descobrisse uma nova forma de se comunicar. Com a comunicação mais fácil e rápida entre pesquisadores e desenvolvedores surgiram várias ferramentas para facilitar a comunicação eletrônica (*E-Mail*: Correio Eletrônico & *Browsers*: Navegadores de páginas) (KRUKOSKI, 2004).

Para Valente (1999, p. 3) “A tecnologia computacional tem mudado a prática de quase todas as atividades, das científicas às de negócio até às empresariais”. Quando passou a ser usada com fins comercial, foi a grande redescoberta do computador. Ficou possível pesquisar, desenvolver, comprar, tudo sem fronteiras, se unir de forma instantânea. Ainda se explora essa área de forma muito tímida diante do potencial que essa comunicação é capaz.

Muitos querem ditar regras ou normas, direcionando um padrão. O importante agora é que todos saibam usar essa tecnologia. Com a facilidade do modo gráfico e o mouse, o mais leigo tem a facilidade de aprender a manusear o computador em muito pouco tempo. (KRUKOSKI, 2004).

## 2. LINGUAGEM NATURAL

Para que se possa compreender melhor como se processa a interação entre o homem e o computador, deve-se procurar entender o processo da comunicação humana.

As pessoas se comunicam para comandar, interrogar, responder, prometer, convencer as outras pessoas. Todas as pessoas quando tem uma idéia e pretendem compartilhá-la com outras, não importando seu objetivo, utilizam-se da comunicação. Comunicar é compartilhar um modelo, torná-lo comum (HUGO, 2004).

Para Hugo (2004) “Uma linguagem é um conjunto de signos e símbolos que permitem um grupo social de se comunicar e facilita o pensamento e as ações dos indivíduos”.

Para que haja a comunicação é necessário um veículo, uma forma de as duas entidades comunicantes compartilharem o mesmo modelo (HUGO, 2004).

### 2.1. Interface homem-computador

Além de um veículo para a comunicação, as duas entidades devem possuir meios de se comunicar em um sentido amplo, uma interface é um dispositivo que serve de limite comum a varias entidades comunicantes, as quais se exprimem em uma linguagem especifica a cada uma. Uma vez que a comunicação esteja estabelecida, a interação ação que é recíproca pode ocorrer entre as entidades (HUGO, 2004).

A tarefa de processar uma linguagem natural permite que os seres humanos se comuniquem com os computadores da forma mais "natural" possível, utilizando a linguagem com a qual mais estão habituados. Elimina-se, desta maneira, a necessidade de adaptação a formas inusitadas de interação, ou mesmo o aprendizado de uma linguagem artificial, cuja sintaxe costuma ser de difícil aprendizado e domínio, a exemplo das linguagens de consulta a bancos de dados (OLIVEIRA, 2000).

Os computadores, como milhares de outras máquinas, são ferramentas que nos ajudam a realizar tarefas de maneira mais eficiente. Carros e aviões, por exemplo, nos permitem viajar com conforto e economizam tempo. Assim como não temos aptidão natural para percorrer longas distâncias, também não estamos preparados para algumas atividades "intelectuais". A superioridade dos computadores em algumas áreas é imensa. Mesmo os computadores mais simples podem fazer cálculos mais rapidamente que qualquer ser humano. Entretanto, como a maioria das máquinas, os computadores são dependentes do ser humano, e precisam ser controlados para obter resultados. Para que sejam úteis, precisam se comunicar com as pessoas. A forma como se comunicam com seus usuários é chamada de interface (O FUTURO, 2000).

No caso da interface homem-computador, a conexão entre as duas entidades se realiza entre a imagem do sistema e os órgãos sensoriais-motores do usuário (HUGO, 2004).

Segundo afirma O Futuro (2000, p.3) “Hoje em dia, uma das mais importantes formas de comunicação com os computadores é o mouse. Usando o mouse podemos apontar para a tela e transmitir ordens precisas”.

Os prognósticos para o futuro revelam a existência de máquinas capazes de compreender a língua natural, sendo capazes, por exemplo, de atender a comandos expressos na linguagem falada por usuários. Ademais, o desenvolvimento de modelos computacionais da língua natural permitirá um maior processamento de informações, visto que a maior parte do conhecimento humano está registrada na forma lingüística. (OLIVEIRA, 2000).

O objetivo da interação homem-computador é, ou deveria ser prover uma interface tão natural quanto possível. De fato, a solução perfeita seria aquela na qual o usuário nem

percebesse a utilização de uma interface. Então, talvez a falta de uma interface seja o nirvana dos usuários de computador. (HUGO, 2004).

Um microfone conectado a um computador pessoal pode possuir várias finalidades, entre elas a de armazenar recados, a de verificar a identidade do locutor para fins de segurança e a de reconhecer palavras na fala. Destas funções, o reconhecimento de voz é a que tem maior potencial para fundamentalmente mudar o modo de interação com o computador: A tecnologia está produzindo uma nova interface homem-computador (HUGO, 2004).

Para Hugo (2004) “As utilizações de reconhecimento de voz e de escrita manual derrubarão as barreiras que teclados e mouses impõem à comunicação natural com o computador”.

## 2.2. Interface em Linguagem natural

Até os anos 70, tinha-se um “grupo” de especialistas em informática para manipular o computador, homens que dominavam a utilização de linguagens especializadas/específicas no uso do computador (HUGO, 2004).

Na década de 80, com o advento e expansão dos computadores e suas redes, o usuário final foi colocado frente a frente com a máquina e seus dados. Desta forma, linguagens mais naturais, bem como meios de acesso mais naturais eram prementes. (HUGO, 2004).

A interface entre computadores e seres humanos é formada principalmente por tela, sistema de som, teclado e mouse. A tela e o som são as formas que o computador tem de se expressar, enquanto o teclado e o mouse são usados para receber instruções. É um processo de mão dupla: digitamos um texto ao mesmo tempo em que os resultados aparecem na tela. Esta comunicação já é uma grande evolução sobre as interfaces mais antigas. Ainda assim, continua precária (O FUTURO, 2000).

Com a evolução da tecnologia, no futuro estaremos trabalhando num ambiente que vai combinar objetos físicos e virtuais. Em nossas mesas os papéis impressos vão conviver com documentos eletrônicos. Post-its virtuais “flutuarão” no ar. E vamos enxergar esse ambiente através de óculos ou lentes especiais, ou algo melhor, que ainda será inventado (O FUTURO, 2000, p. 2).



O crescimento de tamanho e importância das bases de dados, os computadores necessitam ter meios de interpretar a linguagem natural. Desta forma um usuário pode mais facilmente encontrar algum argumento de pesquisa na base de dados, sem se preocupar com sua especificação exata, seja em termos de comandos de busca, seja em palavras a pesquisar, permitindo realizar suas pesquisas através do vocabulário que lhe é conhecido, sendo o computador responsável por oferecer-lhe sinônimos e ajuda direcionada ao argumento (HUGO, 2004).

As pessoas se comunicam por palavras, gestos, olhar, expressões faciais, cheiros e toques. Em algum momento, os computadores terão interfaces completas - e vão interagir suavemente com nossos sentidos (O FUTURO, 2000, p. 1).

Profissionais de informática estimam que a utilização da linguagem natural seja o que há de melhor a oferecer ao usuário em termos de interface. Contudo para haver comunicação é necessário haver diálogo, senão se reduz a uma simples transmissão de informações que não satisfaz ao ser humano (HUGO, 2004).

A utilização de linguagem natural não garante que a interface seja natural. Isto é, fazer com que o usuário possa digitar seus comandos de acordo com seu vocabulário coloquial facilita seu acesso ao computador, porém oferecer-lhe uma interface através da qual ele consiga dar entrada a esta mesma linguagem por voz ou escrita manual (ou ambos) seria mais próximo ao modo comum dele comunicar-se (HUGO, 2004).

Basicamente, as aplicações dos sistemas que tratam à língua natural podem ser divididas em duas classes: aplicações baseadas em texto e aplicações baseadas em diálogos (OLIVEIRA, 2000).

- Exemplos de aplicações baseadas em texto são sistemas que procuram documentos específicos em uma base de dados (exemplo: encontrar livros relevantes em uma biblioteca), tradutores de documentos, e sistemas que resumem textos (exemplo: produzir 3 páginas resumidas de um livro de 100 páginas).
- Com relação às aplicações baseadas em diálogos, podem-se citar as interfaces de linguagem natural para bancos de dados, os sistemas tutores e os sistemas que interpretam e respondem a comandos expressados em linguagem escrita ou falada.

A necessidade de interação natural do homem com o computador aumenta mais, e que estes métodos de entrada/saída naturais vêm sendo vislumbrados para promover a comunicação que acarreta as varias facilidades e vantagens, afirma ainda que a fala seja um dos meios de comunicação mais naturais entre o os seres humanos. Assim, é de suma importância à aquisição e utilização de tecnologias para torná-la realidade nas interfaces de computador. Nas ultimas décadas vários trabalhos foram desenvolvidos, graças ao aumento da capacidade dos computadores e as significativas evoluções nos resultados de pesquisas desenvolvidas contribuiu decisivamente para estes avanços (HUGO, 2004).

Contudo, vale ressaltar que a comunicação humana não ocorre apenas através da fala e sim através de multicanais, formada pelo conjunto dos sentidos. Grandes avanços científicos e tecnológicos ainda estão reservados para estas áreas de pesquisas (HUGO, 2004).

### 3. RECONHECIMENTO DE VOZ

É ponto pacífico que as atuais tecnologias de interface homem/computador para entrada de dados, como teclado e mouse estão fadados a serem substituídas por outras, mais naturais aos seres humanos, principalmente a fala, já que é a principal forma de comunicação e por excelência a melhor interface homem/homem que existe (BATHAGLINI, 1999).

Desde *2001: Uma Odisséia no Espaço*, o homem sonha com o dia em que poderá falar e ser respondido pelo computador, estabelecendo assim com ele um contato completamente diferente de tudo o que se viu até hoje nesse campo. O ano 2001, no entanto, já passou e ainda falta muito para cada um ter o seu próprio *Hall*, tal e qual foram idealizados na ficção científica por Stanley Kubrick. Ainda assim, muito já se avançou nessa área (2001: UMA ODISSÉIA, 1968).

Com o reconhecimento de voz, não só os computadores passaram a ter seu uso facilitado, mas também eletrodomésticos, elevadores, bancos e automóveis. Programas que possibilitam aos computadores reconhecer a voz humana tiveram um avanço notável nos últimos anos. Pode não estar longe o dia em que o computador será capaz de entender sua pergunta, ou pelo menos, de pedir que você a repita. Isso não quer dizer que o computador compreende o que se fala a ele, apenas que ele já é capaz de ouvir e transcrever um texto com um grande índice de acerto. Compreender a fala transcrita envolve uma inteligência que a máquina não tem (VIANA, 1999).

### 3.1. Reconhecimento de Voz

Para Bathaglini (1999, p. 3) “O reconhecimento de voz é o processo no qual se extrai de forma automática a informação necessária do sinal de voz”.

A mágica do reconhecimento de voz ocorre porque a comparação entre a palavra pronunciada pelo locutor e a palavra que se encontra armazenada no banco de dados ocorre em tempo real (VIANA, 1999).

Os programas existentes de reconhecimento de voz ainda não estão prontos para serem usados pela grande maioria, apesar do fato de várias empresas estarem fazendo grandes investimentos nessa área. Esses softwares já avançaram bastante e atingiram o ponto de serem uma bênção para pessoas portadoras de necessidades especiais (VIANA, 1999).

Quando os softwares estiverem prontos para serem usados pelo público maior, eles irão ser incorporados aos sistemas operacionais. Mas, assim como as teclas que ativam o cursor não foram eliminadas com a chegada do mouse, a digitação tão pouco será eliminada (VIANA, 1999).

Uma grande revolução ocorrerá quando os computadores conseguirem realmente entender a voz humana e obedecerem aos comandos verbais. Corporações do mundo digital acreditam que ele será a próxima revolução tecnológica, capaz de causar um impacto similar - se não maior - ao da Internet (VIANA, 1999, p. 2).

A linguagem falada é a forma mais usada de comunicação entre os seres humanos. Devido à capacidade do cérebro de interpretar informações extremamente complexas. Pode-se de forma praticamente inconsciente, captar facilmente em uma mensagem falada várias informações além da transmitida textualmente pelas frases vocalizadas. Reconhecendo assim quem está falando *o agente emissor da mensagem*, sua posição no espaço físico, seus estados emocionais e vários outros dados que podem estar escondidos no tom de voz usado como, por exemplo: ironia, seriedade ou tristeza (BATHAGLINI, 1999).

Porem as máquinas que cercam a todos, e o computador dentre elas, no entanto, não apresentam condições de analisar esse tipo de dados como o cérebro humano o faz. O que de certa forma torna relação homem/máquina um tanto quanto “fria”. Assim, para que se possa realizar a comunicação com as máquinas é necessário que se apertem botões, movam-se alavancas, digitar comandos em teclados, em suma, formas bastante estranhas ao meio normal

de comunicação, a fala. Para tanto, para que se possa trabalhar com as máquinas e isso é essencial no atual estágio de desenvolvimento da civilização, as pessoas devem se adaptar a elas e aprender uma série de métodos diferentes da natureza humana e, portanto, incômodos, principalmente para o cidadão médio e pouco afeito às evoluções tecnológicas (BATHAGLINI, 1999).

Para melhorar a relação homem/máquina, ou a interface entre eles, há décadas cientistas do mundo todo têm pesquisado tecnologias que permitissem o reconhecimento de voz, habilitando assim qualquer pessoa a comunicar-se diretamente e de forma natural com os computadores e os sistemas computacionais que controlam as máquinas que cercam a todos no dia-a-dia. Isso representaria um avanço enorme por vários motivos: Permitiria uma menor curva de aprendizado necessário para o usuário do computador, sem mencionar a velocidade e eficiência com que essa utilização se daria, levando um aumento do desempenho e produção individual e a possibilidade de estar com as mãos livres para se fazer outras coisas enquanto se utiliza o computador ou qualquer outro tipo de equipamento dotado desta tecnologia (BATHAGLINI, 1999).

A tecnologia de reconhecimento de voz possui suas pesquisas direcionadas em varias áreas e setores do mercado mundial, como é o caso de alguns governos mundiais que vem trabalhando em sistemas de reconhecimento de voz para a realização do censo por telefone. O cidadão seria atendido por um computador, responderia verbalmente os dados solicitados e o sistema extrairia automaticamente as informações necessárias das frases e preencheria com elas o formulário do censo, economizando quantias em torno de 2 a 4 bilhões de dólares (BATHAGLINI, 1999).

A empresa americana IBM vem implementando há alguns anos um software utilitário presente em seu sistema operacional OS/2, chamado *VoiceType*, que realiza reconhecimento de voz numa tentativa de permitir a emissão de comandos para o computador e outras coisas interessantes como, por exemplo, o ditado de documentos diretamente para o processador de textos, sem a necessidade de digitação. Os resultados vêm sendo altamente promissores, apesar de ainda não contar com uma grande taxa de acertos (BATHAGLINI, 1999).

Grandes avanços vêm ocorrendo a todo o momento nessa área, permitindo a previsão de ótimos resultados num futuro próximo, emergindo setores da tecnologia ainda não explorados e propiciando um novo viés para os estudos e pesquisas que se farão num futuro não tão distante (BATHAGLINI, 1999).

### 3.2. Funcionamento do Mecanismo de Reconhecimento de Voz

Como afirma Bathaglini (1999, p. 3) “o reconhecimento de voz é o processo no qual se extrai de forma automática a informação necessária do sinal de voz”.

A informação pode ser uma dentre várias as que se apresentam codificadas dentro do sinal de voz emitido, e o sistema reconhecedor pode estar interessado na informação textual da fala para a edição de um texto ou execução de um comando. O sistema também pode estar interessado em algumas palavras chave, como "sim" ou "não", que esteja contida na frase. Ou ainda no timbre pessoal do agente emissor do comando de voz, para a identificação do mesmo como uma medida de segurança (BATHAGLINI, 1999).

O processo de reconhecimento de voz pelo sistema computacional ocorre em três fases distintas: aquisição do sinal de voz, extração de parâmetros e reconhecimento do padrão (VIANA, 1999).

A aquisição do sinal de voz se dá a partir de um dispositivo conversor analógico/digital, que obtém o sinal a ser reconhecido. Em um microcomputador de mesa, esse processo poderia ser o seguinte: um usuário emitiria o sinal de voz em um microfone acoplado a uma placa de som que digitaliza o sinal analógico deixando-o preparado para a próxima fase do processamento (VIANA, 1999).

Na segunda fase, um algoritmo de parametrização, ou seja, um programa desenvolvido para tratar de forma parametrizada o dado de entrada “comando vocal” através de um conjunto de características que descrevem de maneira adequada as propriedades do sinal da voz, extraíndo e representando os parâmetros necessários para a sua utilização (VIANA, 1999).

Após a extração das características do padrão, o reconhecimento do padrão responsabiliza-se pela identificação dos mesmos, isto é, através de comparações sucessivas,

ele verifica a que padrão de referência *conhecida* o padrão de entrada *o qual se deseja reconhecer* se assemelha (BATHAGLINI, 1999).

Nesse ponto se encontram os maiores problemas ainda não solucionados no campo do reconhecimento de voz. Isso porque não se trata somente da simples comparação de dados para se identificar o padrão, já que ele não é fixo. Ocorrem variações nas características da fala que dificultam extremamente o processamento. Supondo que um usuário emita um comando de voz ao computador. Cada vez que ele fizer isso, ocorrerão pequenas diferenças nas formas de onda que compõem o comando, devido à articulação dos órgãos do aparelho fonador. O reconhecimento do comando também pode ser prejudicado pela distância entre o usuário emissor do sinal de voz e do dispositivo de recepção do mesmo ou por algum obstáculo que se interponha entre os dois. O sinal de voz também pode reverberar na parede, ou possuir variações de amplitude como volume menor ou maior, ou ainda ser atrapalhado pela emissão de outro sinal de voz no mesmo recinto por outra pessoa, entre muitíssimos outros problemas que podem ocorrer. A pronúncia do usuário também pode não ser suficientemente correta devido a um sotaque exótico ou algum problema de dicção (BATHAGLINI, 1999).

Além de todos esses problemas, ainda há o da segmentação da fala, o qual tem tomado muito tempo dos cientistas que pesquisam o reconhecimento de voz. Esta imprecisão do limite advém da grande variação dos sinais de voz e a interação mútua entre eles, não se tem precisamente uma forma de limitação dos fonemas (menor unidade da fala), o que dificulta o reconhecimento de fala contínua. Porque ao falar, o som que se emite é algo como: "Abraoprogramadediçãodetexto" e não "Abra-o-programa-de-edição-de-texto" (BATHAGLINI, 1999).

A mesma palavra pronunciada várias vezes podem apresentar diferentes formas de onda devido à articulação dos órgãos do aparelho fonador. As variações nas características da fala: pode haver diferenças acústicas não-lineares no tempo *ritmo*, em frequência *timbre*, e em amplitude *intensidade* (VIANA, 1999).

Com insuficiente uso do conhecimento lingüístico: a fala pode não conter toda a informação lingüística como, por exemplo: erros de Português e sotaque. (VIANA, 1999).

Os seres humanos conseguem facilmente reconhecer todos esses problemas e apreender a informação desejada facilmente e de forma automática de vozes dos mais diversos timbres, amplitudes, com os sotaques mais estranhos e no meio de um emaranhado de ruídos. Porém não se sabe exatamente que mecanismos o cérebro utiliza para que esse processamento seja realizado e, portanto, não se sabe como o processador deve se comportar para obter o mesmo sucesso (BATHAGLINI, 1999).

As restrições acima irão influenciar características como precisão, tipo de aplicação, custo, entre outras. Para contornar algumas restrições foram determinados certos fatores para o reconhecimento (VIANA, 1999):

*Dependência do Locutor:* se o sistema somente reconhece a voz dos locutores para que fosse treinado, tal sistema é dependente do locutor. O sistema é independente do locutor quando é capaz de reconhecer qualquer locutor que não tenha sido treinado.

*Tipo de fala:* podem-se reconhecer palavras isoladas ou fala contínua. No primeiro caso é necessário um período mínimo de silêncio entre as palavras pronunciadas e no segundo esta restrição não é aplicada.

*Tamanho do vocabulário:* o tamanho do vocabulário influencia a precisão do sistema de reconhecimento. Isto ocorre devido a possível ambigüidade das palavras (palavras semelhantes para o algoritmo classificador).

O principal, porém, é que tudo isso criará inúmeras portas para o surgimento de muito mais revoluções, tanto na forma em que as indústrias de tecnologia trabalham, quanto na relação entre usuários e os computadores. Relacionamento com eles será outro completamente diferente, mais direto, mais natural e mais produtivo. Pessoas com necessidades especiais, como os deficientes visuais e os tetraplégicos poderão se utilizar desses sistemas para trabalharem tranquilamente com computadores, sem necessidade da utilização dos atuais meios de entrada e saída de dados. Conversando com o computador, eles serão capazes de fazer o trabalho de praticamente qualquer pessoa de forma tão produtiva quanto. Isso sem contar com o fato de que os computadores não serão as únicas máquinas a utilizar os sistemas de reconhecimento de voz. Carros guiados pela fala, elevadores que se move para um andar específico apenas com uma palavra e luzes que acendem ou apagam ao



som da voz são apenas alguns dos exemplos do que está por vir pela frente para facilitar a vida do ser humano. Em suma, o futuro que se descortina traz inúmeras novidades, mas uma boa parte delas passa necessariamente pelo reconhecimento de voz. (BATHAGLINI, 1999 – VIANA, 1999).

## 4. CONCEITOS GERAIS DA METODOLOGIA UTILIZADA

### 4.1. Método Utilizado

O método utilizado é o dedutivo, pois a forma de apresentação desse trabalho é do aspecto global para o particular. Apresentando de uma forma concisa os conceitos de portadores de necessidades especiais, reconhecimento de voz e linguagem natural, além do estudo e desenvolvimento de uma ferramenta proposta.

### 4.2. Análise Orientada a Objetos

Permitir que os programadores organizassem os programas da mesma forma que a mente humana enxerga os problemas, está é a proposta da orientação a objetos e não como um conjunto de espaços de memória, mas como um conjunto de coisas que fazem parte do problema (MATOS, 2002, p. 20).

Tendo isto como ponto de vista, o programador não será mais obrigado a abstrair dos problemas variáveis e suas respectivas organizações, mas ter o problema como um grande conjunto de objetos. Assim é necessário que se tenha uma modificação na maneira que se enxerga os programas, não sendo permitido neste paradigma, orientar a solução do problema de acordo com as funções identificadas no problema, mas de acordo com o que existe no escopo do problema: Objeto (MATOS, 2002, p. 20).

#### 4.2.1. Objetos

O objeto é um conceito, uma abstração, algo com limites nítidos e significados em relação ao problema em causa, eles facilitam a compreensão do mundo real e oferecem uma base real pra ser implementada no computador (RAMOS, 2003, p. 4).

Um objeto é uma entidade concreta, apesar de ser uma concepção abstrata, tratando-se de um agrupamento de características e ações desta entidade, tendo suas características agrupadas para que se possam identificar outros objetos parecidos (MATOS, 2002, p. 22).

Suas ações ajudam a classificar outras entidades como objetos semelhantes a ele. O jargão utilizado para representar tal característica é o atributo e para representar as ações é o método, sendo assim, um conjunto de atributos e métodos formam um objeto (MATOS, 2002, p. 22).

#### 4.2.2. Encapsulamento

Consiste em separar os aspectos externos de um objeto, dos detalhes internos de implementação. Isto permite que a alteração da implementação de um objeto não afete as aplicações que utilizam. O encapsulamento impede que um programa se torne tão independente que uma pequena modificação possa causar grandes efeitos de propagação, em suma, encapsulamento é um pacote de operações e atributos o qual representa o estado em um tipo de objeto, de tal forma que o estado é acessível ou modificável somente pela interface provida pelo encapsulamento. O encapsulamento também visa o fato de que o mesmo modelo de instrução aparece varias vezes dentro do mesmo sistema (RAMOS, 2003, p. 5).

#### 4.2.3. Classes e Instâncias

Uma classe de objetos descreve um grupo de objetos com propriedades semelhantes, o mesmo comportamento, os mesmos relacionamentos com outros objetos e a mesma semântica (RAMOS, 2003, p. 7).

Uma classe nada mais é que a criação de uma estrutura geral de onde objetos podem ser derivados. Em analogia com jargões de computação, uma classe seria um tipo de dados de onde variáveis pudessem ser criadas. Em orientação a objetos, essas variáveis são instancias de classes, ou seja, objetos (MATOS, 2002, p. 22).

A organização dos programas em classes permite que seja atingido um objetivo ate certo ponto almejado entre todos os paradigmas de programação: a reutilização de código (MATOS, 2002).

#### 4.2.4. Herança

É o compartilhamento de atributos e operações entre classes com base num relacionamento hierárquico. Uma classe pode ser dividida em subclasse mais definidas. Cada subclasse herda todas as propriedades de sua superclasse e acrescenta suas próprias e exclusivas características. As propriedades da superclasse não precisam ser repetidas em cada subclasse (RAMOS, 2003, p. 8).

Para Matos (2002, p.24) “O exemplo da conta de um banco é bastante elucidativo. A classe conta pode ser transformada em uma superclasse para conta poupança e conta especial”.

A herança é uma das principais vantagens dos sistemas baseados em objetos. A capacidade de identificar propriedades comuns às várias classes de uma superclasse comum e de fazê-las herdar as propriedades da sua superclasse pode reduzir substancialmente as repetições nos projetos e programas (RAMOS, 2003, p. 8).

#### 4.2.5. Polimorfismo

É quando a mesma operação pode atuar de modos diversos em classes diferentes. Uma operação é uma ação que um objeto executa. Uma implementação específica de uma operação por uma determinada classe é chamada de método. Pode haver mais de um método para a implementação de um operador baseado em objetos, pois ele é poli mórfico. (RAMOS, 2003. p. 10).

Para Matos (2002, p. 26) “Polimorfismo, enfim, é essa possibilidade de dar a um mesmo método, por exemplo, varias formas, de acordo com o momento em que você decide utilizá-lo”.

Uma linguagem de programação baseada em objetos seleciona automaticamente o método correto para implementar uma operação com base no nome da operação e na classe do objeto que esteja sendo operado. O usuário de uma operação não precisa saber quantos métodos existem para implementar uma determinada operação poli mórfica. Novas classes podem ser adicionadas sem que se modifique o código existente, pois são fornecidos, para cada operação aplicável nas novas classes, os respectivos métodos (RAMOS, 2003, p. 10).

### 4.3. UML (*Unified Modeling Language*)

A UML é uma linguagem criada com o propósito de permitir especificações, visualizações e documentação de artefatos de sistemas. Ou seja, é permitido usar a UML para entender e manipular todos os elementos que direta ou indiretamente influenciam na construção de software. Dessa forma, é possível, também, a avaliação da modelagem do negocio ou de quaisquer outros aspectos não necessariamente computacionais (MATOS, 2002, p. 29).

#### 4.3.1. Diagrama de Caso de Uso

Identifica as funções do sistema e seus respectivos responsáveis, e é tipicamente usado para especificar ou caracterizar a funcionalidade e o comportamento de um sistema de aplicações, interagindo com o mundo exterior (MATOS, 2002, p. 31; RAMOS, 2003, p. 16).

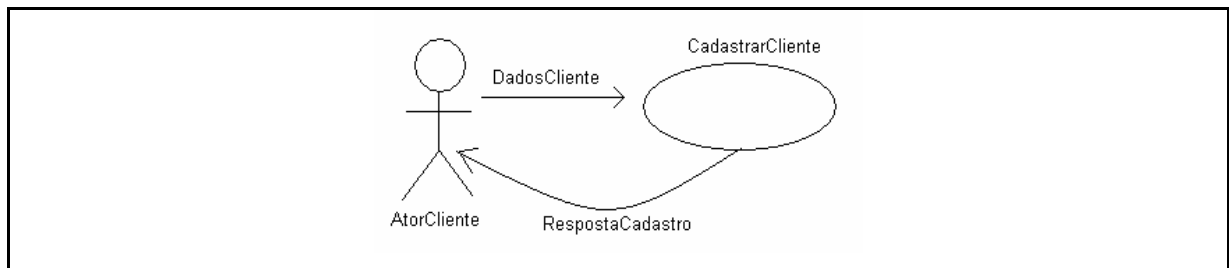


Figura 1 - Diagrama de Casa de Uso  
Fonte: (RAMOS, 2003, p.16).

#### 4.3.2. Diagrama de Classes

Identifica as estruturas mínimas de informação, fazendo uma representação gráfica para descrições genéricas do sistema usado nas fases de análise e projeto de um sistema (MATOS, 2002, p. 31; RAMOS, 2003, p. 15).

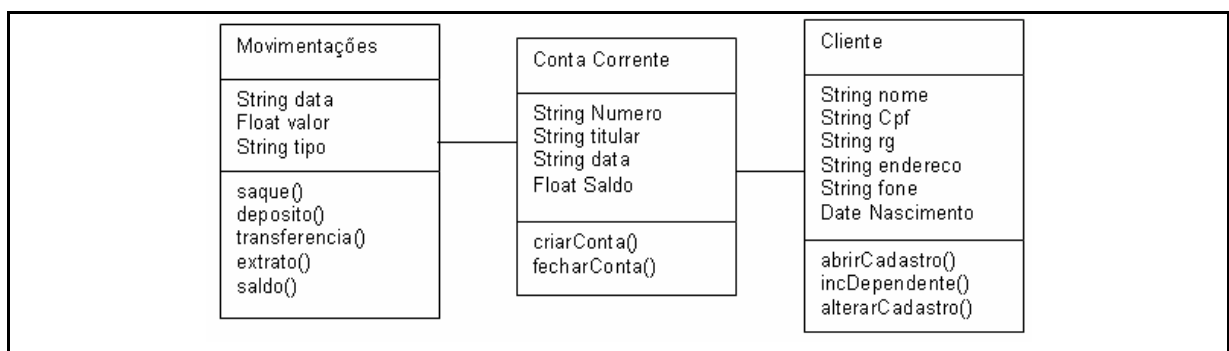


Figura 2 - Diagrama de Classes  
Fonte: (RAMOS, 2003, p. 15).

#### 4.3.3. Diagrama de Interação (seqüência ou colaboração)

Modelar e compreender o comportamento e o dinamismo das classes, mostrando a interação organizada em uma seqüência de tempo entre os objetos participantes de uma operação e as trocas de mensagens entre eles (MATOS, 2002, p. 32; RAMOS, 2003, p. 16).

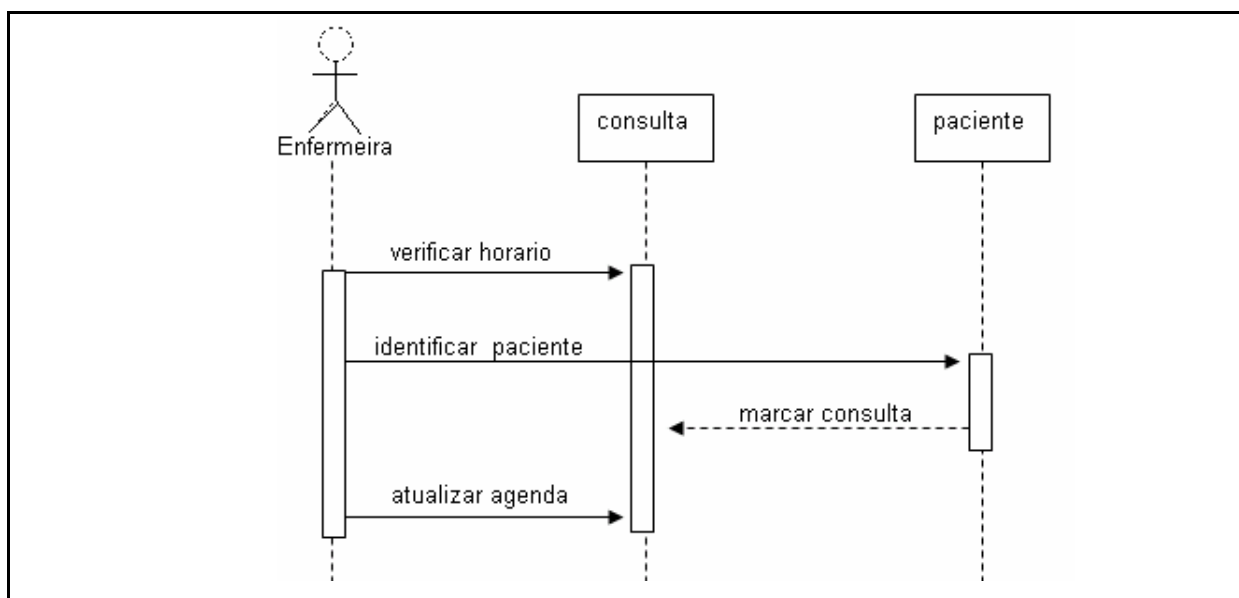


Figura 3 - Diagrama de Interação  
Fonte: (MATOS, 2002, p. 32).

#### 4.3.4. Diagrama de Estado e de Atividade

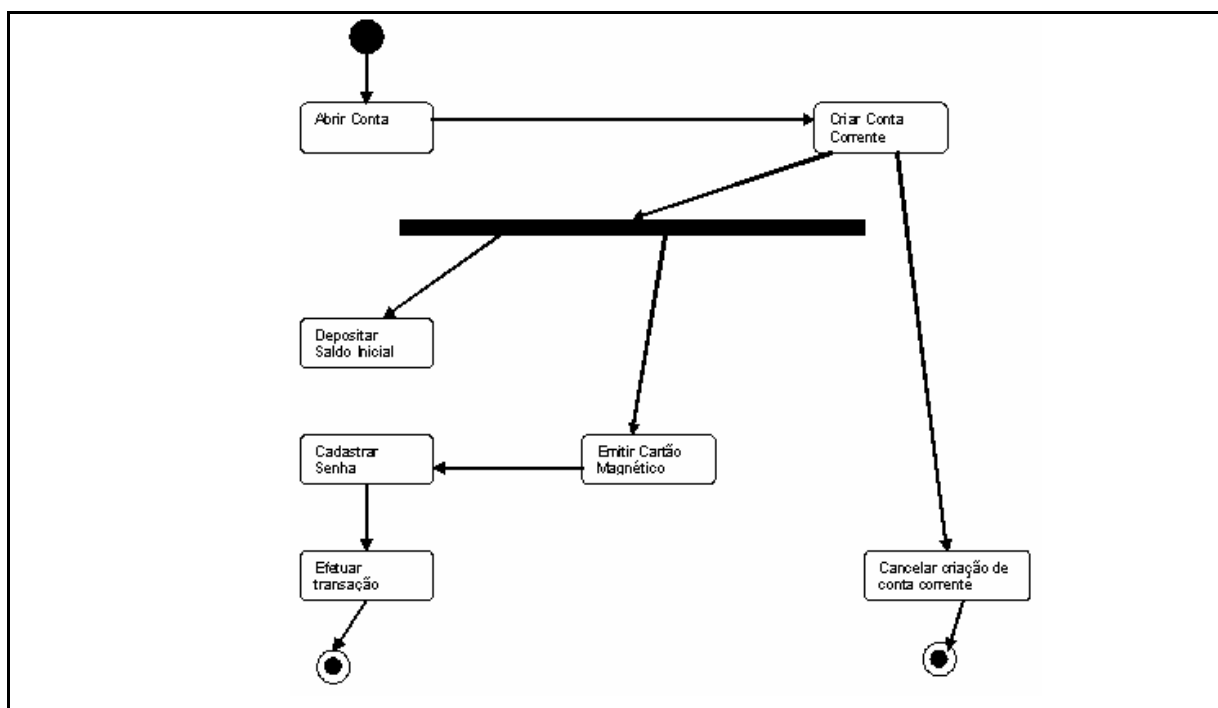


Figura 4 - Diagrama de Atividade  
Fonte: (RAMOS, 2003, p. 18).

É usado para mostrar os estados dos objetos de uma classe. Os eventos do diagrama de estados causam uma transição de um estado para outro e as ações resultam na mudança de estado. Compreende os estados das execuções, permitindo identificar o comportamento das classes. Cada diagrama de estado está associado a uma classe, descrevendo as atividades desempenhadas em uma operação (MATOS, 2002, p. 32; RAMOS, 2003, p. 18).

#### 4.3.5. Diagrama de Componentes e de implantação

Mostra as conexões físicas entre os processadores, dispositivos e a alocação dos processos aos processadores, mostrando a organização do hardware e a ligação do software com os dispositivos físicos (MATOS, 2002, p. 33; RAMOS, 2003, p. 19).

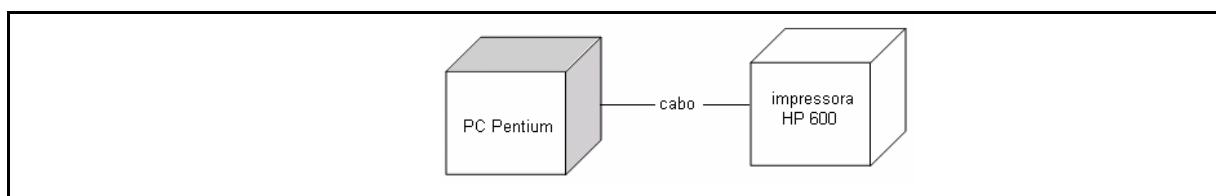


Figura 5 - Diagrama de Componentes e de Implantação  
Fonte: (RAMOS, 2003, p.19).

#### 4.4. Ambiente de Desenvolvimento Delphi

Considerado uma das melhores ferramentas de desenvolvimento de software, atendendo aos objetivos do cliente, permitindo uma aplicação fácil de usar e adequada aos objetivos propostos, mas porque utilizar delphi e não uma outra ferramenta de desenvolvimento? Esta é uma resposta que cabe a cada desenvolvedor escolher (SONNINO, 2001, p. XVI).

Algumas características ajudam ao desenvolvedor escolher o Delphi como ferramenta de desenvolvimento (SONNINO, 2001, p. XVI):

- O delphi é uma ferramenta multiuso: permite desenvolver tanto aplicações comerciais como científicas com a mesma facilidade, sem apresentar problemas de desempenho;
- O delphi gera executável nativo, não sendo interpretado, permitindo assim melhor performance;

- O delphi é orientado a objetos, permitindo programas mais robustos e mais fáceis de depurar;
- O delphi é extensível, podendo-se criar novos componentes ou mesmo expandir sua IDE (*Integrated Development Environment* – Ambiente de Desenvolvimento Integrado), escrevendo código usando o próprio Delphi.
- O delphi integra-se facilmente com o Windows, permitindo elaborar programas que extraem o Máximo desta interface gráfica;
- O delphi é uma ferramenta RAD (*Rapid Application Development*) com a qual, arrastando e soltando componentes, montam-se as janelas da aplicação. Isso permite criar protótipos de aplicações muito rapidamente e, posteriormente “recheiar” o protótipo com o código definitivo, trazendo muita velocidade ao desenvolvimento da aplicação.
- O delphi vem com alguns controles ActiveX previamente instalados e você pode comprar e instalar mais controles ActiveX de terceiros facilmente.

Usando o jargão OLE, um controle ActiveX é um objeto documento composto que é implementado com uma DLL servidora dentro do processo e suporta automação OLE, edição visual e ativação de dentro para fora (inside-out) (CANTÙ, 2002, p. 719).

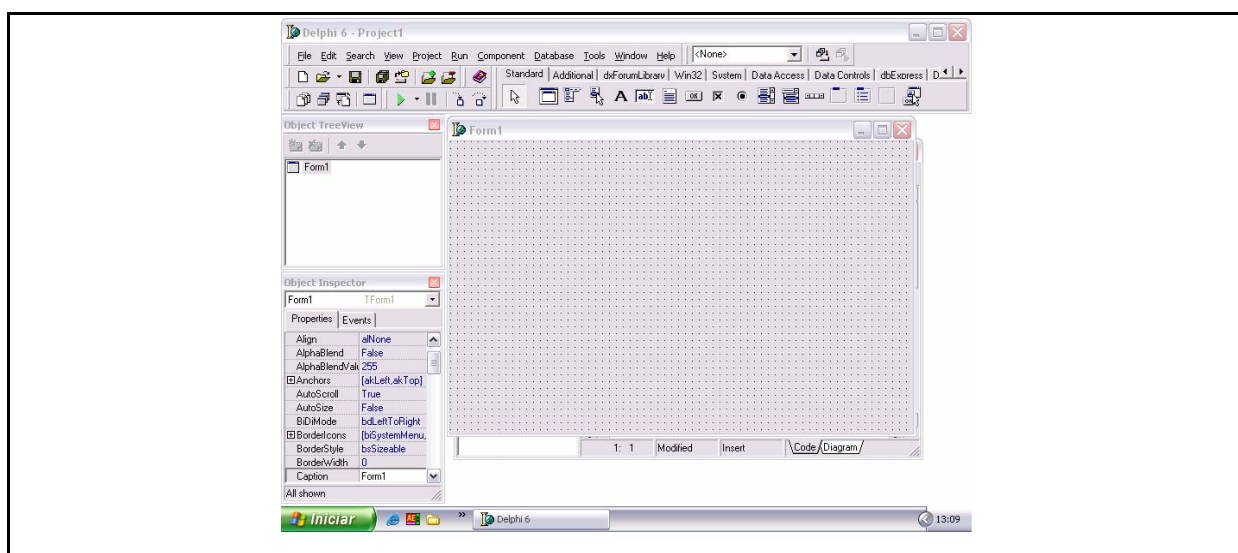


Figura 6 - Ambiente de Desenvolvimento do Delphi

Fonte: Com base na instalação realizada pelo Autor.

No desenvolvimento do sistema FonoMouse utilizou-se um ActiveX (Microsoft Speech SDK 5.1) para realizar o processo de reconhecimento de voz e sendo o delphi um ambiente de desenvolvimento que se integra facilmente com os ActiveX (SONNINO, 2001, p. XVI). Esta afirmação e o conhecimento prévio deste ambiente de desenvolvimento foi o que



levou a escolha do Delphi como ferramenta de desenvolvimento utilizado na construção do sistema.

#### 4.5. Microsoft Speech

O Microsoft Speech SDK é um conjunto de ferramentas, disponibilizada em um controle ActiveX, que adicionado ao código fonte de um aplicativo durante a sua construção, por intermédio de uma linguagem de programação, prove as funcionalidades de reconhecimento de voz e síntese de voz ao software desenvolvido (MICROSOFT, 2004).

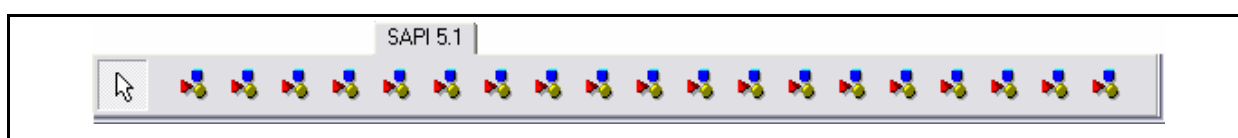


Figura 7 - Paleta de Componentes Speech SDK instalado no Delphi  
Fonte: Com base na instalação realiza pelo autor.

#### 4.6. Modelagem de Dados do FonoMouse

Este tópico apresenta os diagramas utilizados no desenvolvimento do sistema, tendo como ponto de vista a melhor visualização dos componentes utilizados na construção do sistema, bem como parte da própria documentação do FonoMouse.

##### 4.6.1. Diagrama de Caso de Uso do FonoMouse

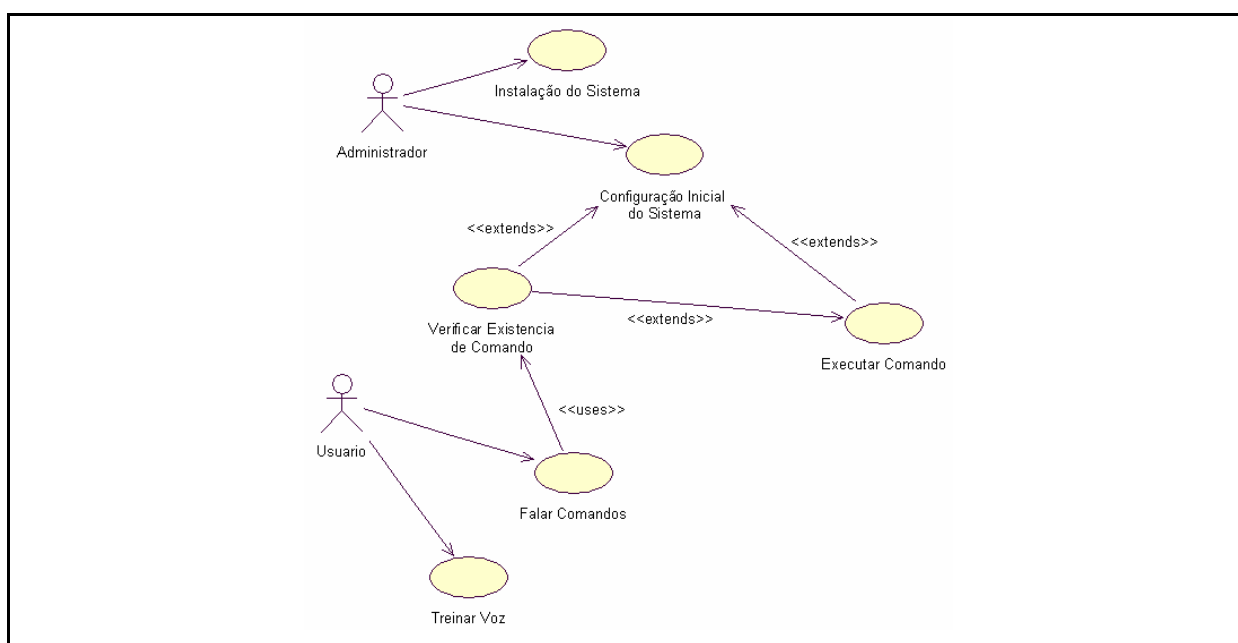


Figura 8 - Diagrama de Caso de Uso do FonoMouse  
Fonte: Com Base na UML desenvolvida pelo Autor.

O administrador do sistema é solicitado a instalar o Sistema, assim como a sua configuração inicial, feito está primeira parte o portador de necessidades especiais, aqui chamado de usuário, poderá então realizar o treinamento da voz. Após o treinamento o sistema já está pronto para ser utilizado, bastando agora o usuário falar os comandos, onde é verificada a existência de cada comando, caso haja, é então executado o comando solicitado, como mostra a Figura 8.

#### 4.6.2. Diagrama de Classes do FonoMouse

O sistema é composto de quatro classes principais, como mostra a Figura 9, onde a classe *Ferramentas* se encarrega de gerar um arquivo de dados da configuração do sistema, para que numa posterior reinicialização do sistema ele mantenha as configurações personalizadas do usuário. A classe *Grade* fica responsável pela criação e posterior destruição de uma grade criada para auxiliar na movimentação do mouse. A classe *MaquinaReco* fica por sua vez responsável pela criação do objeto de reconhecimento de voz, assim como a verificação constante da existência do comando falado, e por fim a classe *Principal* que é a responsável pela obtenção do sinal de voz, que é passado à classe *MaquinaReco*, e caso tenha uma resposta afirmativa de que o comando exista, ela então realiza a ação solicitada.

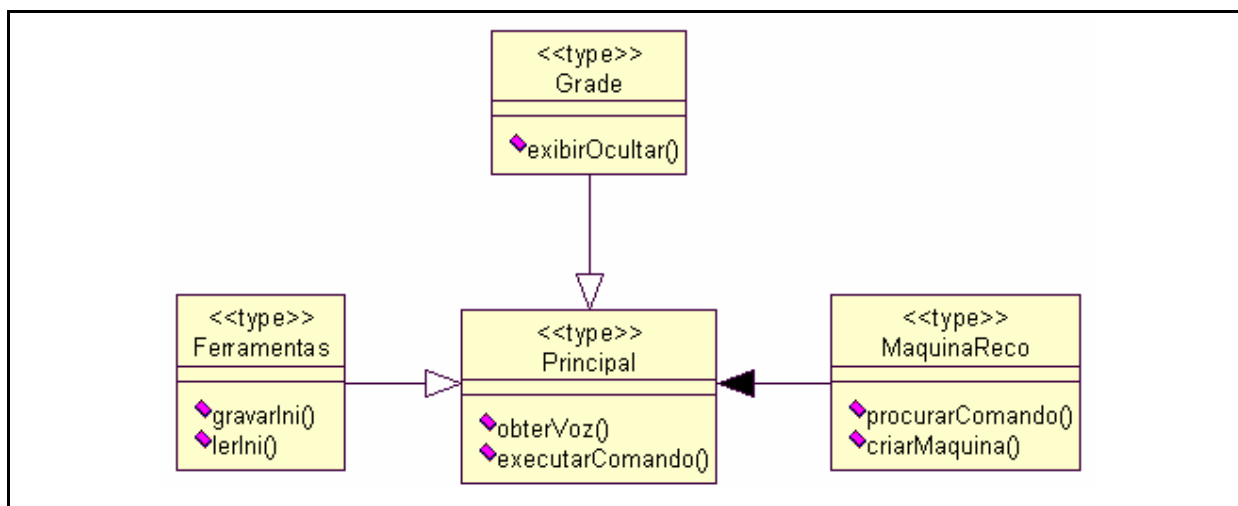


Figura 9 - Diagrama de Classe do FonoMouse

Fonte: Com Base na UML desenvolvida Pelo Autor.

#### 4.6.3. Diagrama de Interação do FonoMouse

O usuário inicia o sistema, que por sua vez verifica o arquivo de dados de configuração, realiza todas as configurações ali armazenadas, exceto a criação da grade de auxilio que é feito logo após a criação da máquina de reconhecimento dos comandos, como

mostra a Figura 10. Feito estas configurações o sistema devolve ao usuário a tela principal onde fica aguardando os comandos. Ao receber um comando o sistema procura a existência do mesmo, e o executa caso ele exista, após a execução do comando, o sistema volta a ficar aguardando o próximo comando, caso o próximo comando seja o de finalizar o sistema, o sistema então grava as configurações existentes e o finaliza.

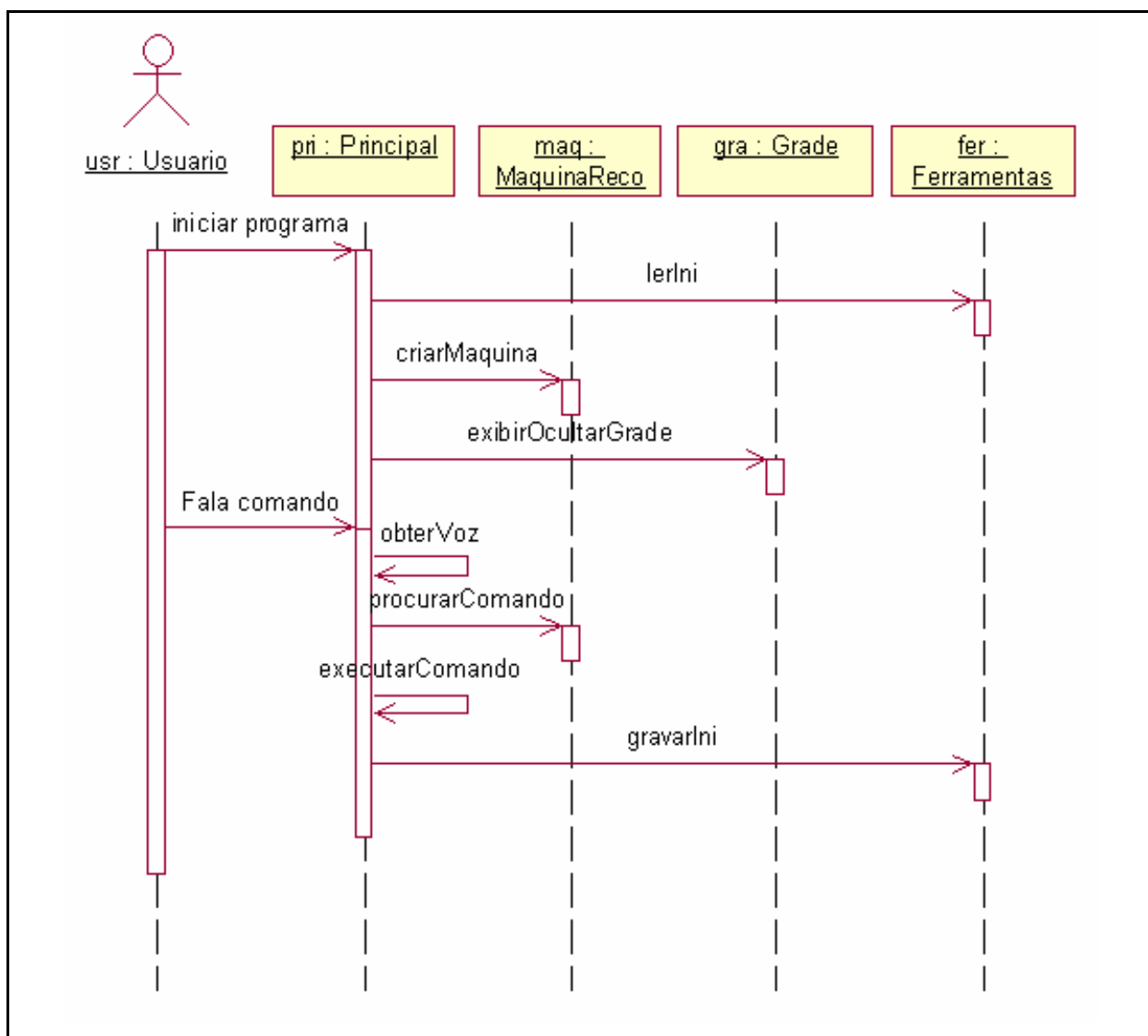


Figura 10 - Diagrama de Interação do FonoMouse  
Fonte: Com Base na UML desenvolvida Pelo Autor.

#### 4.6.4. Diagrama de Estado e de Atividade do FonoMouse

Durante o estado de espera do sistema, ou seja, aguardando algum comando, qualquer ruído de som que ocorra é passado à função de procurar comando que verifica a existência do comando, caso o comando exista ele é executado, caso não, ele volta ao seu estado de espera.

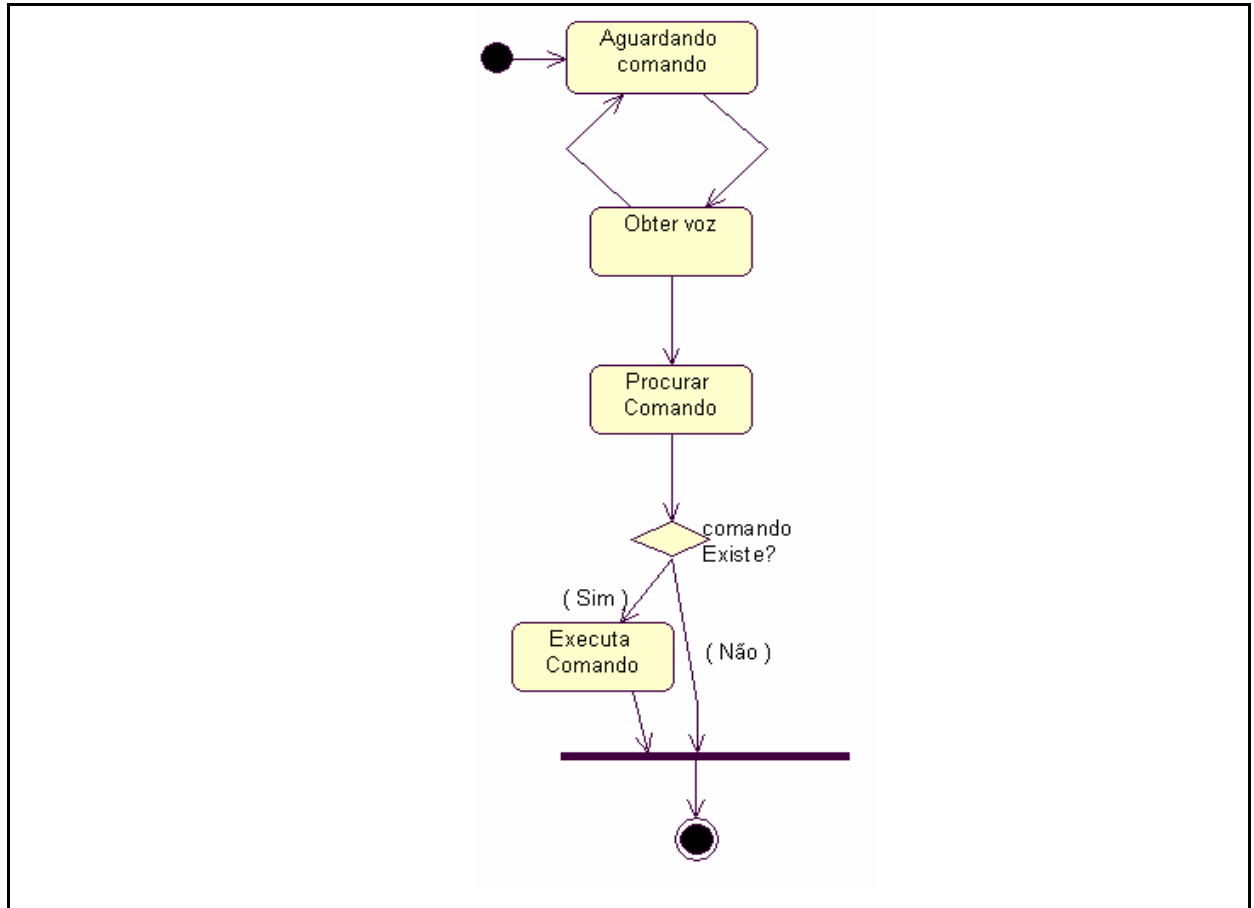


Figura 11 - Diagrama de Estado e de Atividade do FonoMouse  
 Fonte: Com Base na UML desenvolvida Pelo Autor.

#### 4.6.5. Diagrama de Componente e de implantação do FonoMouse

O som do comando é obtido através de um microfone ligado a uma placa de som, que por sua vez esta interligada ao sistema através do processador.

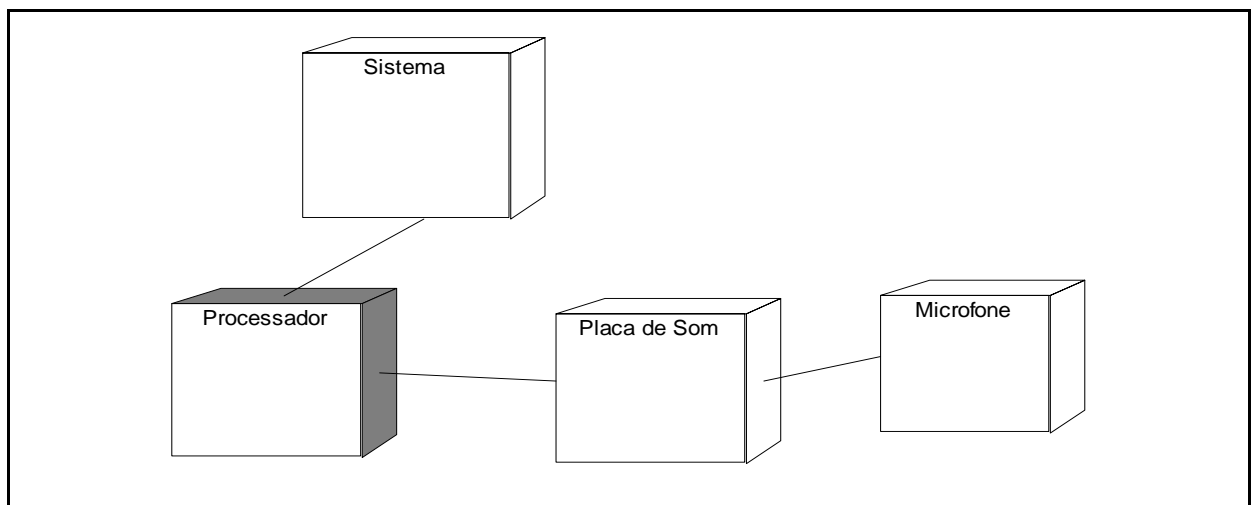


Figura 12 - Diagrama de Componente e de Implantação do FonoMouse  
 Fonte: Com Base na UML desenvolvida Pelo Autor.

## 5. FUNCIONALIDADES DO SISTEMA

### 5.1. Tela de carregamento

Na tela da Figura 13 é apresentada a logomarca do sistema, em quanto em segundo plano é carregado para a memória o sistema.



Figura 13 - Tela de Carregamento do FonoMouse  
Fonte: Com Base no desenvolvimento do Software proposto.

### 5.2. Tela Principal

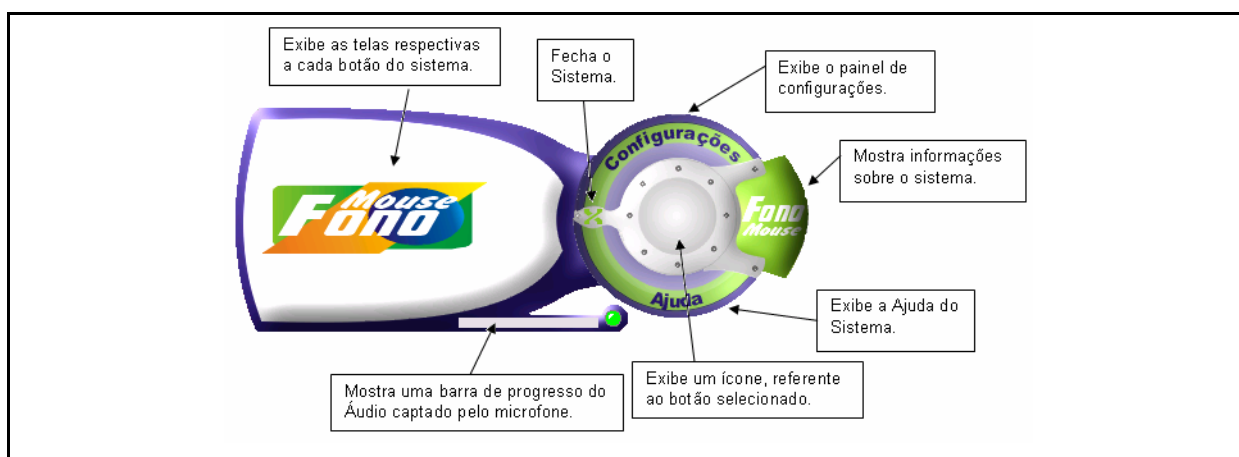


Figura 14 - Tela Principal do Sistema  
Fonte: Com Base no desenvolvimento do Software proposto.

A Figura 14 apresenta o início do sistema onde se dá todas as opções de navegação pelo programa, é apresentado ainda os quatros botões que são os principais meios de acesso ao sistema, sendo eles: Configurações, Ajuda, Fechar e FonoMouse.

### 5.3. Painel de Configurações

Exibe os botões de configurações de mouse, grade, voz e microfone os quais permitem realizar as configurações necessárias ao bom funcionamento do sistema, como mostra a Figura 15.

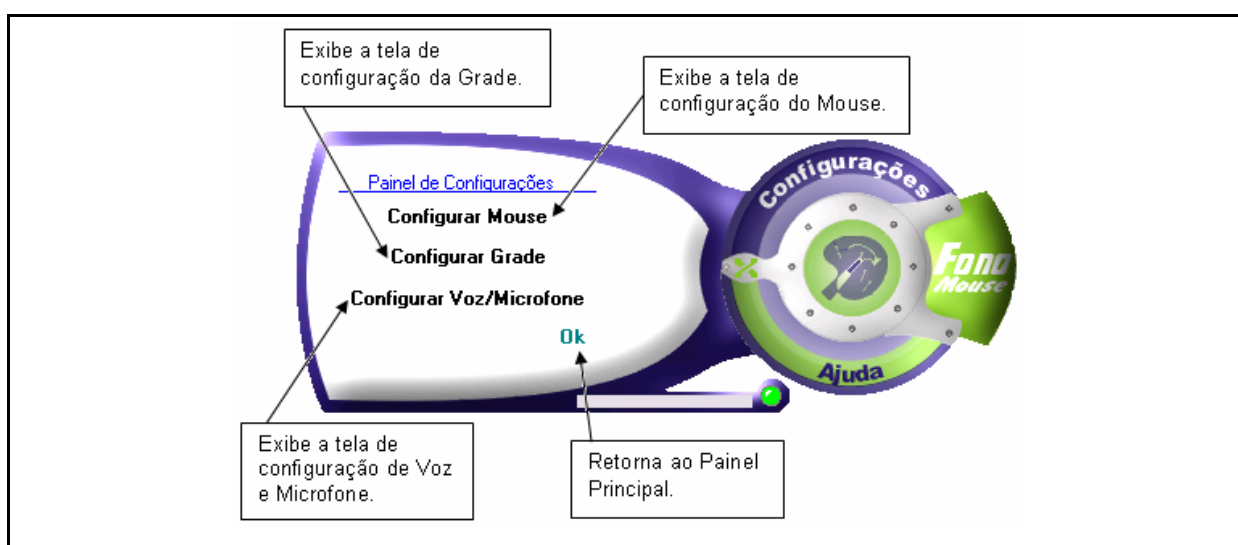


Figura 15 - Painel de Configurações

Fonte: Com Base no desenvolvimento do Software proposto.

#### 5.3.1. Configurar Mouse

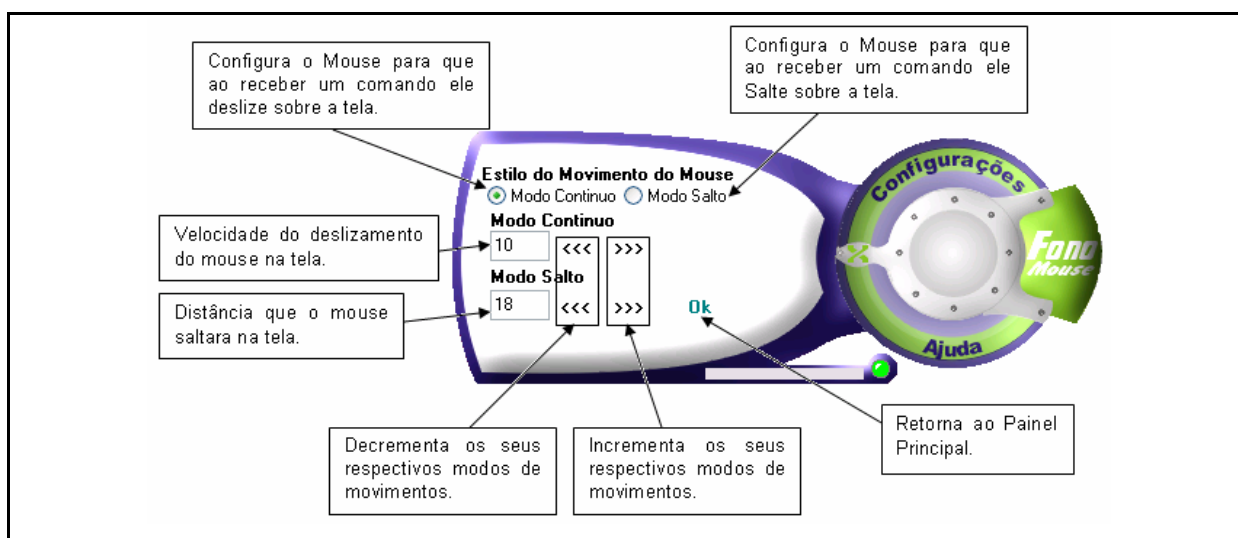


Figura 16 - Configurações do Mouse

Fonte: Com Base no desenvolvimento do Software proposto.

Exibe as opções disponíveis no sistema para que se possa configurar o mouse da maneira que mais se adaptar o usuário, demonstrado na Figura 16.

### 5.3.2. Configurar Grade

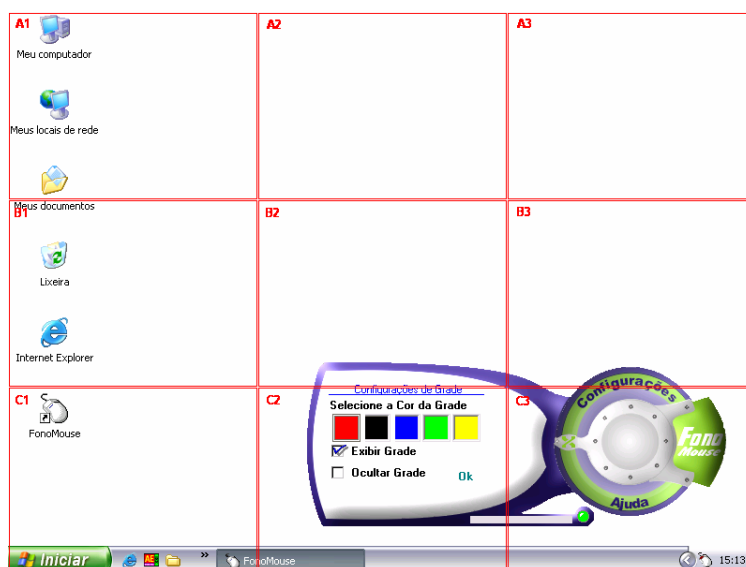


Figura 17 - Exibição da Grade

Fonte: Com Base no desenvolvimento do Software proposto.

Exibe as opções necessárias para a configuração da grade, que serve como auxílio na movimentação do mouse na tela, onde a tela é dividida em uma grade com nove células, como mostra a Figura 17, caso seja escolhida a opção de exibir grade, ainda nesta tela é possível selecionar a cor da grade, demonstrado na Figura 18.

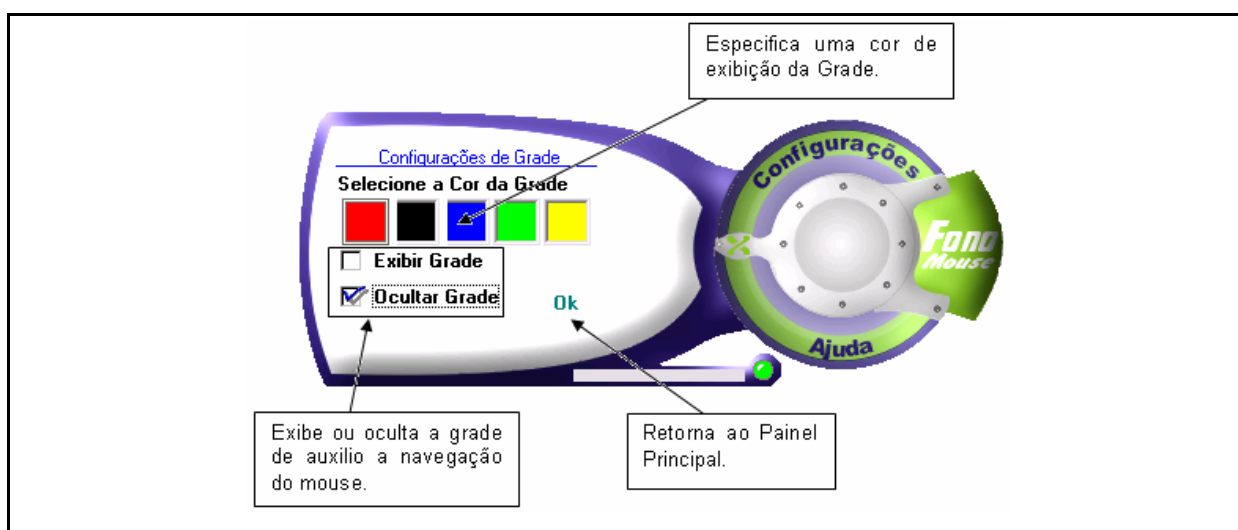


Figura 18 - Configurações de Grade

Fonte: Com Base no desenvolvimento do Software proposto.

### 5.3.3. Configurar voz e microfone

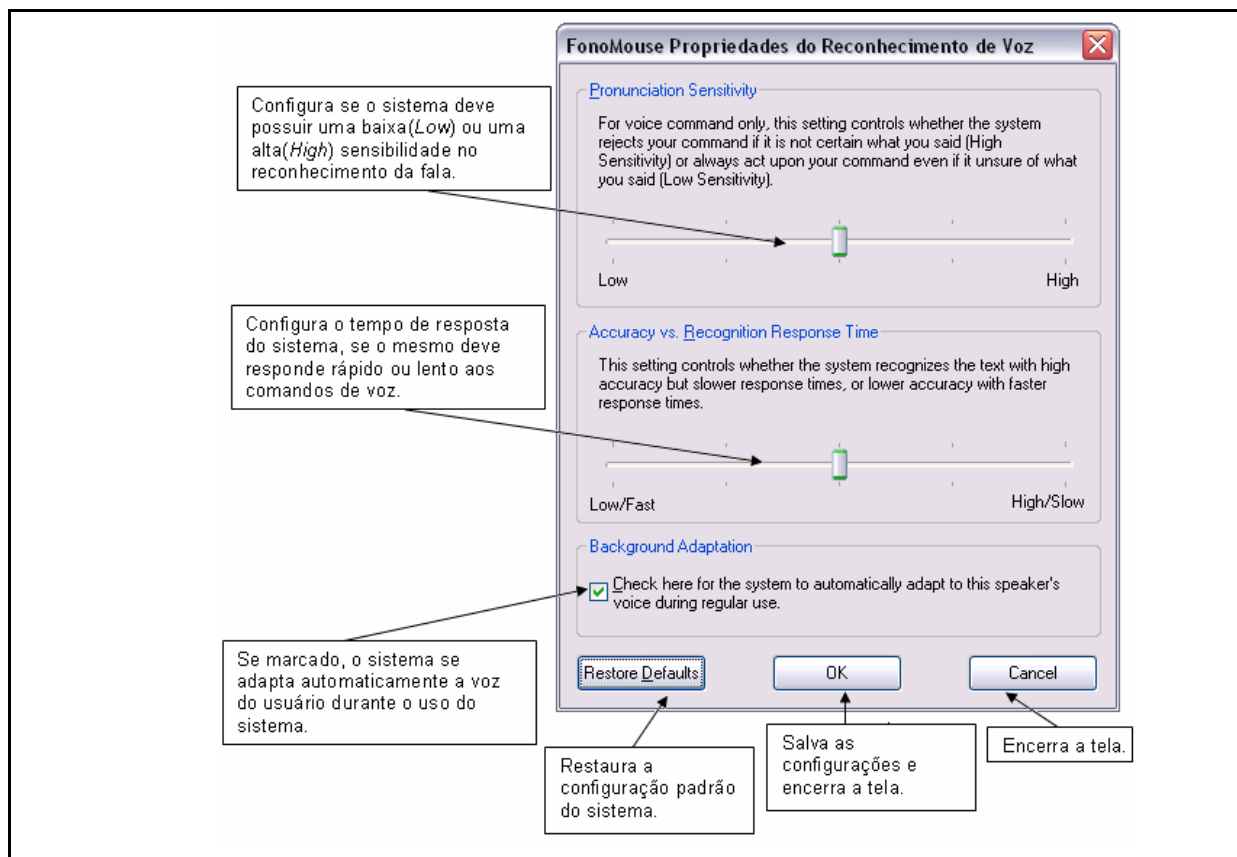


Figura 19 - Configurações de Voz

Fonte: Com Base no desenvolvimento do Software proposto.

Exibe as opções de configuração de voz e microfone, para que o sistema possa obter um melhor desempenho no reconhecimento da fala do usuário, como mostra a Figura 20.

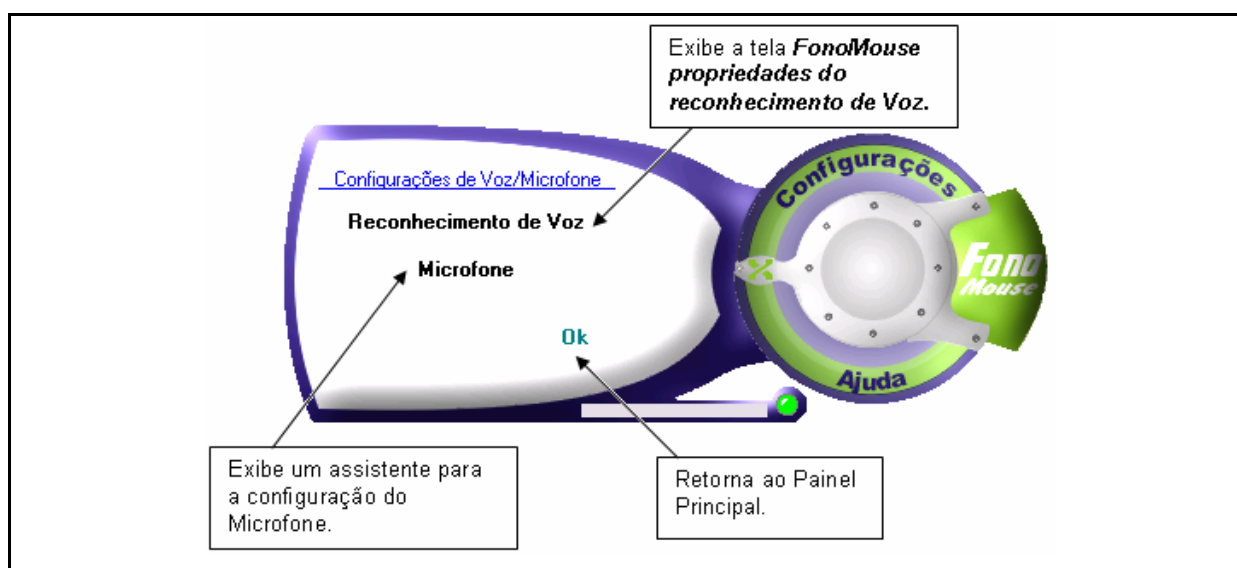


Figura 20 - Configuração de Voz e Microfone

Fonte: Com Base no desenvolvimento do Software proposto.



Ao ser acionado o botão *Reconhecimento de Voz* é exibida a tela de propriedades do reconhecimento de voz demonstrado na Figura 19, para configurar a sensibilidade do reconhecimento da voz, já o botão *Microfone* exibe um assistente de configuração do microfone demonstrado na Figura 21, neste assistente é configura o ruído do ambiente para se obter uma melhor performance de reconhecimento.



Figura 21 - Assistente de Configuração do Microfone  
Fonte: Com Base no desenvolvimento do Software proposto.

#### 5.4. Ajuda

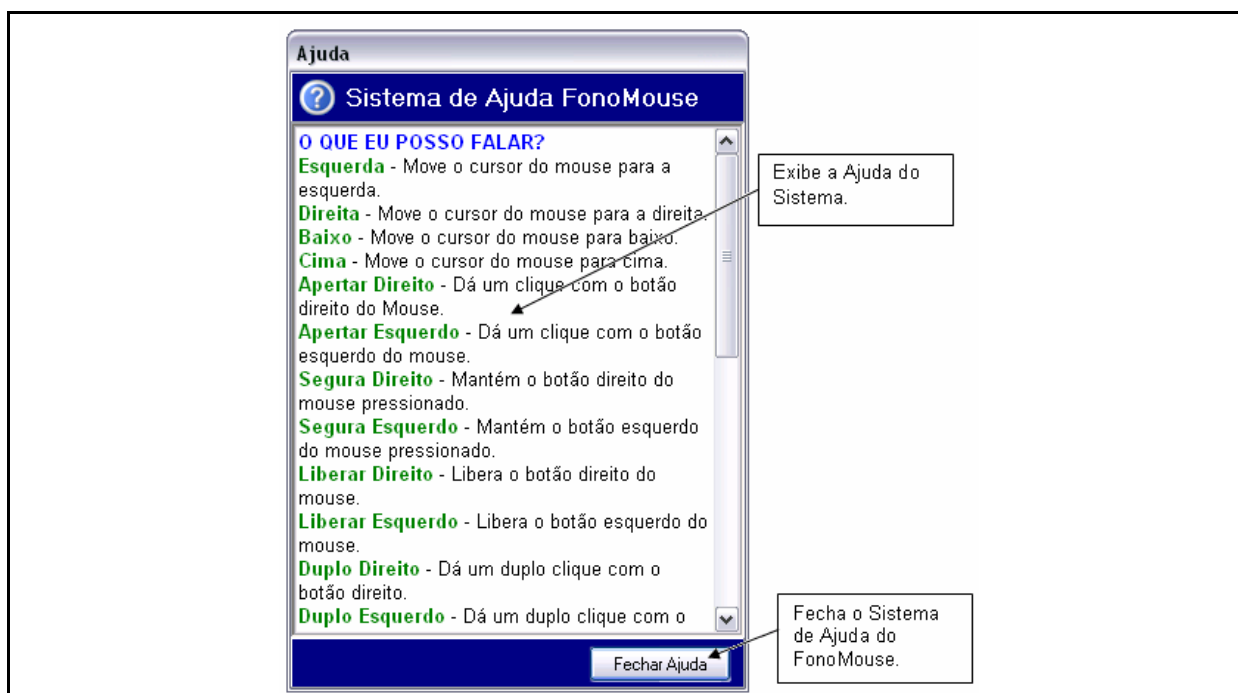


Figura 22 - Sistema de Ajuda do FonoMouse  
Fonte: Com Base no desenvolvimento do Software proposto.

Exibe a tela Sistema de Ajuda do FonoMouse, onde é apresentado à ajuda do sistema, como mostra a Figura 22, especificando quais os comandos que o usuário pode falar e uma descrição das principais configurações do sistema.

### 5.5. Sobre o FonoMouse

Exibe informações sobre o Sistema FonoMouse, mostrando também a versão do sistema, como é apresentado na Figura 23.

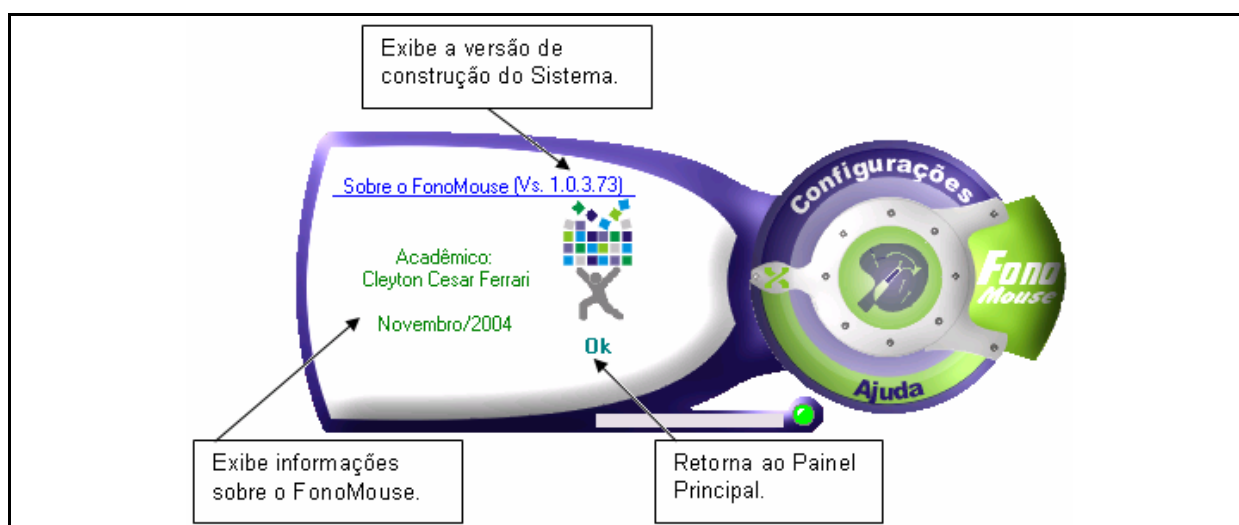


Figura 23 - Sobre o FonoMouse

Fonte: Com Base no desenvolvimento do Software proposto.

### 5.6. Requisitos de instalação do FonoMouse

Os requisitos para a instalação recomendadas estão demonstrada nos itens seguintes, a instalação em ambientes não recomendados podem causar erros de execução do sistema.

#### *Sistemas Operacionais Suportados*

- Windows XP Professional ou Home Edition;
- Microsoft Windows 2000 todas as versões;
- Microsoft Windows Millennium Edition;
- Microsoft Windows 98 todas as versões;

#### *Hardwares Necessários*

- Intel Pentium 400Mhz, superior ou equivalente;
- 128 megabytes (MB);

- 90 megabytes (MB) de espaço disponível em disco rígido;
- Placa de som compatível com a versão do Windows instalada, com um conector de saída de microfone ou uma porta USB disponível para entrada de microfone USB;
- Um microfone de alta qualidade e com filtro de ruídos. A qualidade do reconhecimento de voz esta significativamente relacionada ao microfone utilizado.

## CONCLUSÃO

A principal meta deste trabalho é desenvolver uma ferramenta que vem auxiliar portadores de necessidades especiais. O modelo proposto (FonoMouse) proporcionou que este objetivo fosse alcançado. Os estudos realizados proporcionaram também a descoberta de dois campos de aplicação para reconhecimento de voz. O primeiro é a conversão fala-texto, sua aplicação é mais imediata, como ditar cartas e e-mails. O segundo é fala-comando (utilizada no sistema), com aplicações muito mais amplas. Uma palavra, devidamente traduzida pelo processador, desencadeia uma ação instrumentalizada, como movimentar o mouse na tela, ou simular o pressionamento de um de seus botões.

No entanto, muito ainda falta avançar no desenvolvimento de uma máquina de reconhecimento de voz, utilizando o português do Brasil, haja vista a dificuldade encontrada no treinamento das palavras em português durante o desenvolvimento do sistema, pois a máquina de reconhecimento de voz utilizada traz consigo somente o dicionário em inglês.

Sendo assim, abrem-se portas para realização de pesquisas nessa área, já que segundo a Microsoft (2004) “o reconhecimento de voz é o novo nicho de mercado”, não é a toa que dados encontrados no próprio site da Microsoft apontam que a empresa investiu 45 milhões de dólares na companhia Lermount & Hauspie, da Bélgica, especializada em programas de reconhecimento de voz. E estão previstos investimentos pela Intel em cerca de 30 milhões de dólares para desenvolver tecnologias de reconhecimento de voz para integrá-las em seus servidores e “*desktops*” (VIANA, 1999).

A área de pesquisa sobre reconhecimento de voz é muito ampla, abrindo assim uma lacuna para uma futura pesquisa de integração do sistema desenvolvido com outros sistemas, como por exemplo, o sistema de fala-texto, que possibilitaria ao usuário além da movimentação do mouse a edição de documentos escritos.

## REFERÊNCIAS BIBLIOGRÁFICAS

BATHAGLINI, Maicon Guerra. **Reconhecimento de Voz**. Centro de Tecnologia da Universidade Federal de Santa Maria - RS, [1999?]. Disponível em: <http://www.inf.ufsm.br>. Acessado em: 20 Abr. 2004.

BORTOLON, Rodrigo; MARTIGNAGO, Abraão; MARCEL, Angelo. **Grupo de Pesquisas em Informática**: Bacharelado em Sistemas de Informação. [S.l.]: Sociedade Paranaense de Ensino e Informática - Faculdades SPEI, [2004?].

BRASIL. Decreto-Lei, de 20 de dezembro de 1999. Regulamenta a Lei no 7.853, de 24 de outubro de 1989, dispõe sobre a Política Nacional para a Integração da Pessoa Portadora de Deficiência, consolida as normas de proteção, e dá outras providências. Brasília, 1999.

BRITO, Christina May Moran de. **Causas Mais Comum de Amputações nos Membros Superiores e Inferiores**. São Paulo: [s.n.], 2004. DMR - Divisão de Medicina de Reabilitação do Hospital das Clínicas da Faculdade de Medicina da Universidade Federal de São Paulo. Disponível em: [http://www.saudeparavoce.com.br/telefisiatria/aulasdidaticas/Z89.9\\_1](http://www.saudeparavoce.com.br/telefisiatria/aulasdidaticas/Z89.9_1). Acessado em: 02 Set. 2004.

CANTÙ, Marcos. **Dominando o Delphi 6 “A Bíblia”**. Trad. João Eduardo Nóbrega Tortello. São Paulo: MAKRON Books, 2002.

HÜBNER, Jomi Fred – **Interface em Linguagem Natural**. Blumenau: Universidade Regional de Blumenau, 2000.

HUGO, Marcelo. **Interfaces Naturais**. Santa Catarina: Universidade Federal de Santa Catarina, [2004?]. Disponível em: [http://www.eps.ufsc.br/disserta/hugo/cap\\_2/cp2\\_hug.htm](http://www.eps.ufsc.br/disserta/hugo/cap_2/cp2_hug.htm). Acessado em: 13 Set. 2004.

LIMA, Teófilo Lourenço de. **Manual básico para elaboração de monografia**. Canoas: ULBRA, 1999.

KRUKOSKI, Eduardo Bruno da Costa. **Entendendo o Computador**. Santa Catarina: Universidade Federal de Santa Catarina UFSC, 2004. Disponível em:

[http://www.mtm.ufsc.br/~krukoski/Entendendo\\_o\\_Computador.html](http://www.mtm.ufsc.br/~krukoski/Entendendo_o_Computador.html). Acessado em: 09 Set. 2004.

MATOS, Alexandre Veloso de. **UML (Unified Modeling Language): Prático e Descomplicado**. 3.ed. São Paulo: Editora Érica, 2002.

MICROSOFT. **Microsoft Speech SDK home page**. [S.l.]: Microsoft, 2004. Disponível em: <http://www.microsoft.com/speech>. Acessado em: 08 Jul. 2004.

MIRANDA, Carlos Roberto. **Oportunidade de Trabalho para portadores de deficiência**. [S.l.:s.n], 2004. Disponível em: <http://www.saudeetrabalho.com.br/download/oportunidades-miranda.doc>. Acessado em: 09 Set. 2004.

NIKOS, Drakos - **Free Speech Journal - Phoneme Probality Estimation with Dynamic Sparsely Connected Artificial Neural Network** – [S.l.:s.n], [2004?]. Disponível em: <http://cslu.cse.ogi.edu/fsj/html/>. - Acessado em: 06 Maio 2004.

O FUTURO da Interface. **Jornal do Brasil**, [S.l.], 7 Set. 2000. Disponível em: <http://www.kischi.com.br/artigos/artigo08.htm>. Acessado em: 13 Set. 2004.

OLIVEIRA, Fabio Abreu Dias de. **Processamento de linguagem natural: princípios básicos e a implementação de um analisador sintático de sentenças da língua portuguesa**. 2000. Dissertação – Instituto de Informática, Universidade Federal, Rio Grande do Sul.

PAULA, Mauricio Braga de – **Reconhecimento de Palavras Faladas Utilizando Redes Neurais Artificiais**; 2000, (Dissertação apresentada ao Instituto de Física e Matemática) - Universidade Federal de Pelotas, Pelotas.

PROJETO Mortrix. **Universidade Federal do Rio de Janeiro: Núcleo de Computação Eletrônica**. Rio de Janeiro: [s.n.], [2002?]. Disponível em: <http://intervox.nce.ufrj.br/motrix/tetraplegia.htm>. Acessado em: 02 Set. 2004.

PUC Minas. **Cartilha da inclusão dos direitos das pessoas com deficiência**. Belo Horizonte: PUC Minas, 2000.

RAMOS, José Marcio Benite. **Sistemas Orientados a Objetos I e II**. Apostila do Curso de pós-graduação em Desenvolvimento para Web. Porto Velho: FATEC-RO, 2003.

REIS, Mauricio Caruzo. **Reconhecimento de Voz: Quase uma Realidade**. [S.l.]: Eletronicaria Noticias, [2001?]. Disponível em: <http://www.letronet.com.br/psist/pnotic/pnoticea/ea01/pnot0201/pnottec.htm>. Acessado em: 23 Abr. 2004.

SANTOS, Alexandro dos; PELIZZONI, Jorge Marques. **Introdução ao processamento de línguas naturais**. [2003?]. LABIC-ICMC-USP (Projeto SIAE-USP). São Carlos/SP. Disponível em: <http://labic.icmc.usp.br/didatico/pln>. Acessado em: 13 Set. 2004.

SONNINO, Bruno. **Desenvolvendo Aplicações com Delphi 6**. São Paulo: MAKRON Books, 2001.

VALENTE, José Armando. **O computador na sociedade do conhecimento**. Campinas: UNICAMP/NIED, 1999.

VIANA, Santos Pedroso; ROSA, Lucas Halberstadt da; SILVA, Gabriela Jacques da. **Reconhecimento de Voz**. Centro de Tecnologia da Universidade Federal de Santa Maria - RS, [1999?]. Disponível em: <http://www.inf.ufsm.br>. Acessado em: 20 Abr. 2004.

2001: UMA ODISSÉIA no Espaço. Direção: Stanley Kubrick. Inglaterra: MGM, 1968. 1 filme (141 min), son., color., 35mm.



## MARCAS REGISTRADAS

ActiveX, Microsoft Speech SDK, Windows XP Professional, Microsoft XP Home Edition, Microsoft Windows 2000, Microsoft Windows Millennium Edition, Microsoft Windows 98, Windows 95: *Microsoft Company*.

OS/2, VoiceType: *IBM Company*.

Intel Pentium: *Intel Company*.

HP 600: Hewlett-Packard Development Company.

Delphi: Borland Company.

O autor declara estar usando as marcas citadas nesta obra apenas para fins acadêmicos, em benefício do dono da marca, sem a intenção de infringir as regras de seu uso.

## ANEXOS