

Paulo Sérgio Santoro

Portal de Voz com foco em Governo Eletrônico

Instituto de Pesquisas Tecnológicas do Estado de São Paulo

Paulo Sérgio Santoro

Interfaces Audíveis - Portal de Voz com foco em Governo Eletrônico

São Paulo

2006

Paulo Sérgio Santoro

Interfaces Audíveis - Portal de Voz com foco em Governo Eletrônico

Dissertação apresentada ao Instituto de Pesquisas Tecnológicas do Estado de São Paulo - IPT, para obtenção do título de mestre em Engenharia de Computação.

Área de concentração: Engenharia de Software

Orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Lucia Vilela Leite Filgueiras

São Paulo

2006

Ficha Catalográfica  
Elaborada pelo Departamento de Acervo e Informação Tecnológica – DAIT  
do Instituto de Pesquisas Tecnológicas do Estado de São Paulo - IPT

S237i Santoro, Paulo Sérgio  
Interfaces audíveis – portal de voz com foco em governo eletrônico. / Paulo Sérgio Santoro. São Paulo, 2006.  
127p.

Dissertação (Mestrado em Engenharia de Computação) - Instituto de Pesquisas Tecnológicas do Estado de São Paulo. Área de concentração: Engenharia de Software.

Orientador: Profa. Dra. Lucia Vilela Leite Filgueiras

1. Portal de voz 2. Governo eletrônico (e-gov) 3. São Paulo (estado) 4. Interface audível 5. Inclusão social 6. Tese I. Instituto de Pesquisas Tecnológicas do Estado de São Paulo. Coordenadoria de Ensino Tecnológico II. Título

07-10

CDU 004.52(043)

## Dedicatória

Dedico esse trabalho aos meus pais, as minhas duas filhas e aos meus familiares. Também o faço, a todas as pessoas que acreditam que as maiores barreiras para se atingir um sonho, estão dentro de nós mesmos, e que basta apenas um pouco de coragem para vencê-las.

## Agradecimentos

Inicialmente gostaria de agradecer à Companhia de Processamento de Dados do Estado de São Paulo - PRODESP pelo apoio e pela confiança a mim depositada. A possibilidade de participar no processo de seleção para o programa de bolsas de estudo para o curso de mestrado do IPT, demonstra o quanto a empresa incorporou os programas de inclusão social e digital propostas pelo poder público. A isenção e a transparência no processo de avaliação, livre de preconceitos, propiciou que um portador de necessidades especiais como eu, pudesse concorrer a essa bolsa de igual para igual, com os demais funcionários da empresa. Parabéns e agradeço a alta cúpula da empresa e em especial ao departamento de recursos humanos da PRODESP, que com esse tipo de atitude prova que não está apenas cumprindo a lei, e sim, aplicando de forma integral os conceitos de responsabilidade social por ela recomendado.

Aos meus familiares que me acompanharam do início ao final dessa jornada, agradeço do fundo do meu coração. Vocês foram fundamentais para que eu pudesse atingir esse título de mestre, me ajudando nos momentos mais difíceis e me incentivando nos momentos de maior insegurança. Houveram algumas situações que pensei que não iria chegar lá, mas entre eles sempre havia alguma palavra de apoio e conforto que me fizeram não desistir. Também gostaria de agradecer aos meus amigos da PRODESP, que sempre me deram o suporte necessário para que eu pudesse chegar a esse objetivo. Não irei citar o nome de todos, pois poderia correr o risco de esquecer de um deles. Porém, eles sabem o quanto estou grato pela ajuda tão importante que eles me forneceram. Outro grupo de pessoas que me auxiliaram nesse processo foram os profissionais do IPT, desde os membros do corpo docente, até a equipe da secretaria, que sempre ofereceram os subsídios necessários e informações preciosas para que eu pudesse chegar ao final desse curso com sucesso. Agradeço, em especial a Ester, que sempre se colocou a minha disposição para ajudar na formatação e padronização do trabalho final.

Por fim, agradeço a prof. doutora Lúcia que não se limitou apenas em me orientar durante o processo de concepção do trabalho, fornecendo todo o apoio e suporte técnico necessário para que eu pudesse chegar até aqui. O seu conhecimento na área de IHC foi fundamental para que pudéssemos eliminar falhas de interpretação nos conceitos das novas tecnologias abordadas ao longo do mesmo. Como também, o seu nível de comprometimento e de

profissionalismo é incrível, sempre atenta aos mínimos detalhes e procurando obter consistência nas informações fornecidas no trabalho. Em resumo, você foi muito mais do que uma orientadora, foi uma amiga. Hoje reconheço que todo o esforço e dedicação valeram a pena.

## RESUMO

A inclusão digital é uma iniciativa que tem a intenção de permitir aos diversos segmentos da sociedade o acesso às vantagens que a tecnologia da informação oferece aos seus usuários. No Brasil, órgãos públicos e instituições não governamentais estão empenhadas no sentido de tornar este projeto uma realidade nacional. Porém, a sua implementação não é simples, mesmo porque, além de proporcionar aos cidadãos acesso aos sistemas computacionais, esse processo faz parte de um projeto muito mais ambicioso, que é o da inclusão social.

As interfaces audíveis podem ser um grande agente facilitador para atingir esse objetivo, pois elas utilizam o som como a principal forma de interação entre o ser humano e o computador. As pessoas idosas, analfabetas e portadoras de necessidades especiais, permanentes ou temporárias, podem ser beneficiadas com a concretização desse projeto, que utiliza o meio de comunicação mais natural do ser humano, que é a fala.

Portanto, o principal objetivo deste estudo é propor um portal de voz com foco em governo eletrônico para o Estado de São Paulo, que sirva como exemplo da capacidade das interfaces audíveis contribuírem de forma efetiva para a inclusão social. A proposta secundária é fornecer uma visão geral das interfaces audíveis de fala e de não fala na atualidade, bem como abordar os principais aspectos da tecnologia utilizada para o desenvolvimento deste tipo de aplicação e destacar alguns fatores de acessibilidade. Deseja-se que esta visão geral contribua para a aceleração das atividades de desenvolvimento de aplicações audíveis com o intuito de auxiliar em futuros estudos, principalmente nos serviços de governo eletrônico no Brasil.

Palavras Chave: Acessibilidade ; Portal de Voz; Interfaces Audíveis; Modelagem Interativa



## ABSTRACT

The digital inclusion is a program made to allow different segments of society to access the advantages offered by IT (Information Technology) for the users.

In Brazil public organs and non-governmental institutions are working together to present this as a national reality. However this implementation is not so simple, it involves the accesses to the computational system and much more; insert citizens in digital inclusion.

The audible interfaces can be a great agent to make this goal easier because it uses sound as the main interaction between machines and the human being. Aged, illiterate, and disabled people (permanently or not) can take advantage of this project that uses the most common way of communication: the ability to speak.

So, the main purpose of this study is to present an e-government voice portal for the state of Sao Paulo as an example on how audible interfaces can be useful in the digital inclusion. The second proposal is to provide a general view on existing speech and non-speech audible interfaces as well show the main aspects of the actual used technology for development and accessibility applications in this platform.

This general view intends to be a helpful tool to speedy up audible application developments activities for future projects, especially those in Brazilian e-government services.

Key words: Accessibility; Voice Portal; audible interfaces; Interactive modeling

## **Lista de abreviaturas e siglas**

<b>ANN</b>	<i>Artificial Neural Networks</i>
<b>ASR</b>	<i>Automatic Speech Recognition</i>
<b>CGI:</b>	<i>Common Gateway Interface</i>
<b>CRM</b>	<i>Customer Relationship Management</i>
<b>CTI</b>	<i>Computer Telephony Integration</i>
<b>DTMF:</b>	<i>Dual Tone Multi-Frequency</i>
<b>DTW</b>	<i>Dynamic Time Warping</i>
<b>e-MAG</b>	<i>Modelo de Acessibilidade de Governo Eletrônico</i>
<b>ERP</b>	<i>Enterprise Resource Planning</i>
<b>HMM</b>	<i>Hidden Markov Models</i>
<b>IHC</b>	<i>Interface Humano Computador</i>
<b>IPVA</b>	<i>Imposto sobre Propriedade de Veículos Automotores</i>
<b>IVR</b>	<i>Interactive Voice Response</i>
<b>MoLIC</b>	<i>Modelling Language for Interaction as Conversation</i>
<b>PDA</b>	<i>Personal Digital Assistants</i>
<b>PVG</b>	<i>Portal de Voz Governamental</i>
<b>SALT</b>	<i>Speech Language Application Tags</i>
<b>SDK</b>	<i>Software Developers Kit</i>
<b>SRGS</b>	<i>Speech Recognition Grammar Specification Markup Language</i>
<b>SSML</b>	<i>Speech Synthesis Markup Language</i>
<b>TTS</b>	<i>Text-To-Speech</i>
<b>URA</b>	<i>Unidade de Resposta Audível</i>
<b>URI</b>	<i>Universal Resource Identifiers</i>
<b>VXML ou VoiceXML</b>	<i>Voice Extensible Markup Language</i>
<b>XML</b>	<i>Extensible Markup Language</i>

## SUMÁRIO

1	INTRODUÇÃO.....	8
1.1	Motivação.....	10
1.2	Objetivo.....	12
1.3	Contribuições do trabalho.....	13
1.4	Metodologia.....	13
1.5	Organização do trabalho.....	14
2	INTERFACES AUDÍVEIS.....	16
2.1	Conceitos Gerais.....	17
2.1.1	Som.....	17
2.1.1.1	Altura ou Tom.....	18
2.1.1.2	Intensidade.....	18
2.1.1.3	Timbre.....	18
2.1.2	Audição.....	18
2.1.3	O som como forma de interface.....	20
2.2	Utilização de Interfaces Audíveis.....	21
2.3	Heurísticas para as interfaces audíveis e Portais de Voz.....	22
2.4	Interfaces audíveis de fala.....	25
2.4.1	Sistema de Reconhecimento de Voz.....	26
2.4.2	Unidade de Resposta Audível.....	27
2.4.3	Sistema de Reprodução da Fala.....	29
2.5	Interfaces audíveis de não fala.....	31
2.5.1	Ícones Audíveis.....	32
2.5.2	Earcons.....	32
2.5.3	Técnicas de Exibição.....	32
2.5.3.1	Técnica sintática.....	33
2.5.3.2	Técnica semântica.....	33
2.5.3.3	Técnica léxica.....	34
2.5.4	Avaliações Audíveis.....	35
2.5.4.1	Espaço audível em 2D.....	35
2.5.4.2	Espaço audível em 3D.....	35
3	MODELAGEM PARA SISTEMAS INTERATIVOS.....	37
3.1	Semiótica.....	37
3.2	MoLIC – Modeling Language for Interaction as Conversation.....	38
3.3	Cenários.....	38
3.3.1	Signos.....	39
3.3.2	Perguntas.....	41
3.4	Modelo de Tarefa.....	41
3.4.1	Diagrama Hierárquico de Metas.....	42
3.4.2	Estruturas de Tarefas.....	43
3.4.3	Tipos de Tarefas.....	43
3.4.4	Pré-condições.....	44
3.4.5	Re-uso de Tarefas e Metas.....	44
3.4.6	Representação dos Signos no Modelo de Tarefas.....	44
3.4.7	Prevenção e Tratamento de Erro.....	45

3.4.8	Estrutura Textual Completa do Modelo de Tarefas.....	46
3.5	Modelo de Interação .....	47
3.5.1	Modelo Diagramático de Interações.....	49
3.5.2	Especificação Textual.....	49
3.6	Considerações finais .....	50
4	PORTAL DE VOZ COM FOCO EM GOVERNO ELETRÔNICO.....	51
4.1	Migração do Sítio de Governo Eletrônico para um Portal de Voz.....	52
4.1.1	Quanto ao Padrão .....	52
4.1.2	Quanto à Funcionalidade .....	53
4.1.3	Quanto ao Desenho.....	53
4.1.4	Quanto ao uso .....	54
4.2	Problemas e Soluções Durante o Processo de Migração .....	54
4.3	Linguagens e Padrões da Tecnologia de Portais de Voz.....	57
4.3.1	VoiceXML ou VXML .....	58
4.3.2	SALT.....	60
4.3.3	Comparação entre VoiceXML e SALT .....	63
4.3.3.1	Plataforma atual de desenvolvimento.....	63
4.3.3.2	Fornecedores da Solução .....	64
4.3.3.3	A aceitação dos profissionais .....	64
4.3.3.4	Característica da aplicação a ser desenvolvida.....	65
4.4	Principais Plataformas de Desenvolvimento desse Tipo de Tecnologia no Brasil.....	65
4.4.1	Nuance V-Builder .....	65
4.4.2	WebSphere Voice Server.....	66
4.4.3	OracleAS 10G .....	67
4.4.4	Microsoft .NET.....	68
5	IMPLEMENTAÇÃO DA MODELAGEM DO IPVA PARA O PORTAL DE VOZ GOVERNAMENTAL .....	70
5.1	Modelagem através da MoLIC para o IPVA.....	70
5.1.1	Cenários Identificados para o IPVA.....	71
5.1.1.1	Perguntas Exploradas nos Cenários.....	87
5.1.2	Modelo de Tarefas do IPVA.....	90
5.1.2.1	Metas do Usuário.....	90
5.1.3	Modelo de Interação para o IPVA.....	98
5.1.3.1	Modelo Diagramático de Interação .....	98
5.1.3.2	Estrutura Textual .....	100
5.2	Usabilidade no Portal de Voz Governamental .....	102
5.2.1	Eficácia.....	102
5.2.2	Eficiência.....	103
5.2.3	Satisfação.....	104
5.3	Contribuição da MoLIC no Portal de Voz Governamental .....	105
6	CONCLUSÃO.....	109
6.1	Dificuldades Identificadas .....	111
6.2	Contribuição da MoLIC.....	112
6.3	Implementação do Portal de Voz Governamental.....	114
6.4	Trabalhos Futuros.....	115
	REFERÊNCIAS .....	117

## 1. INTRODUÇÃO

Os sistemas computacionais, para que possam ser utilizados de forma adequada, necessitam que haja entre o usuário e aplicação, um conjunto de procedimentos e técnicas de interação eficaz. Para tanto, durante o processo de desenvolvimento de um sistema, o profissional de informática deve lançar mão de todos os recursos disponíveis para facilitar a comunicação entre essas duas entidades.

As interfaces computacionais fazem exatamente este papel, pois são os elos de ligação entre o ser humano e o computador. Sendo assim, agem como um elemento de fundamental importância no processo de interação, determinando a aceitação de uma aplicação, ou não, pelo usuário final. Elas podem ser divididas, quanto à forma de comunicação, em dois grandes grupos, compostos pelas interfaces visuais e pelas interfaces audíveis. As primeiras podem ser representadas pelas interfaces gráficas, que utilizam imagens estáticas ou dinâmicas como forma de interação. Já as segundas podem ser representadas pelas interfaces audíveis de fala e de não fala, que utilizam a voz e os sons naturais ou artificiais com este mesmo propósito.

As interfaces audíveis, durante muito tempo, ficaram em segundo plano durante o processo de desenvolvimento de um novo sistema, ao serem comparadas com as interfaces visuais. Este fato pode ser atribuído, em parte, a alguns fatores circunstanciais, tais como o alto custo de implantação das aplicações audíveis, no que se refere à tecnologia e equipamentos necessários para sua execução de forma satisfatória, bem como a grande aceitação das interfaces gráficas pelos usuários, já que tem o poder de transmitir uma grande quantidade de informação, através do uso de imagens.

Porém, esse quadro veio se alterando gradualmente, até que na atualidade, isso passou a não ser mais uma verdade absoluta. A gama de aplicações que contemplam o som como meio de comunicação tem crescido de forma significativa. Os portais de voz são um grande exemplo desta mudança e já podem ser encontrados com relativa facilidade em várias aplicações comerciais. Esse crescimento constante na utilização dessas interfaces audíveis e o interesse do público em geral, ou seja, usuários e desenvolvedores, têm aumentado a cada dia. O que significa que a responsabilidade dos profissionais que desenvolvem este tipo de sistema é cada vez maior. Contudo, observa-se que não há muita divulgação dos métodos e processos disponíveis para o seu desenvolvimento, estando eles mais restritos às comunidades

específicas, como o meio acadêmico, científico e militar.

Cabe ressaltar, que apesar dessas dificuldades, as interfaces audíveis têm alcançado um lugar de destaque e de grande importância nos sistemas computacionais em desenvolvimento na atualidade. Com o objetivo de reforçar essas afirmações, são fornecidos a seguir alguns dados obtidos em um artigo publicado no sítio da SUCESU-ES, citado por Contantine (2003). Baseado em algumas pesquisas como a do Evans Group, realizada com 250 clientes da Nuance, uma das principais empresas que fornecem soluções de automação com o uso da voz humana, o artigo apontou que 83% dos usuários preferiam utilizar o reconhecimento de voz a discar e que 74% preferiam atendimentos dirigidos com utilização de reconhecimento de voz, em vez de falar com um atendente.

Esses portais, por permitirem um fácil acesso às informações, são um excelente instrumento para atender às necessidades básicas de um público heterogêneo, como cidadãos de um determinado Município, Estado ou Nação. Desta forma, os serviços oferecidos pelos governos poderão ser bastante democratizados e úteis para a população em geral.

A principal proposta desse estudo é a de empregar essa nova tecnologia, utilizando a *Modelling Language for Interaction as Conversation* (MoLIC) que é uma linguagem de modelagem baseada nos conceitos da Engenharia Semiótica, para demonstrar a autenticação e a navegação de um usuário dentro de um portal governamental, o qual fornece um conjunto de serviços oferecidos das diversas Secretarias de Estado, através do reconhecimento da voz humana. Neste portal é possível obter várias informações de interesse dos contribuintes. Dada a abrangência dos serviços prestados pelos governos, o foco é o de modelar uma aplicação de governo eletrônico da Secretaria dos Negócios da Fazenda do Estado de São Paulo, mais especificamente, na execução de um acesso típico, que um contribuinte faria no sistema do IPVA (Imposto sobre Propriedade de Veículos Automotores).

A Engenharia Semiótica define o processo de interação entre o ser humano e o computador, como sendo uma forma particular de comunicação humana intermediada por sistemas computacionais. Dessa maneira, a interação é vista como um diálogo entre o usuário e o projetista, mediada pelo aplicativo. Ela propõe uma análise minuciosa das necessidades do usuário ao utilizar o aplicativo, no que se refere as metas, objetivos e as atividades a serem executadas dentro do sistema. Essa teoria está baseada na premissa de que se o usuário

consegue aprender como o sistema funciona e porque o sistema foi construído dessa maneira, ele se torna capaz de utilizar o aplicativo sem encontrar maiores dificuldades.

A MoLIC por ser uma linguagem baseada nos princípios da Engenharia Semiótica, implementa esses conceitos no seu processo de modelagem e é composta por três fases durante o seu desenvolvimento. A primeira, contém os cenários que fornecem uma visão geral dos processos comunicativos e pode ser complementada com perguntas criadas pelo projetista, com o objetivo de sanar algumas dúvidas existentes nos diálogos. Nesta fase, também são definidos os signos que serão utilizados nas interações entre o sistema e o usuário. Já na segunda fase, o foco é o de se descobrir e detalhar as metas e os objetivos do usuário ao utilizar o sistema. Por fim, na terceira fase são descritos minuciosamente todos os processos interativos, identificados nas fases anteriores.

## 1.1 Motivação

Apesar da grande gama de serviços disponíveis e da facilidade de acesso às informações que a Internet oferece aos seus usuários, infelizmente, apenas uma pequena parcela da população brasileira tem acesso aos recursos e vantagens oferecidos por este tipo de tecnologia, seja por fatores de origem cultural, social, ou mesmo, por uma opção pessoal. De fato, na realidade, somente apenas alguns privilegiados desfrutam dos benefícios que a informática proporciona aos seus usuários. Os programas de inclusão digital têm trabalhado no sentido de diminuir esta grande lacuna, porém encontram várias dificuldades, no que diz respeito à freqüente falta de verba para os programas sociais, à heterogeneidade do público alvo, aos problemas sócio-culturais e às dimensões continentais do país que, entre outros fatores, fazem com que o processo se torne lento e de difícil implementação.

A solução para o problema em questão não é trivial, mesmo porque envolve uma enorme variedade de fatores humanos e técnicos, isto é, como fazer para democratizar o uso do computador e levar essa tecnologia à maioria das pessoas. As respostas para essas indagações foram surgindo ao longo do tempo e somente na atualidade se tornaram possíveis.

O que há algum tempo atrás não passava de mera ficção científica, hoje já é uma realidade. A tecnologia de reconhecimento da fala humana trouxe à luz um antigo sonho do ser humano, que era o de conversar com a máquina. Cabe ressaltar que a forma de comunicação mais natural do ser humano é através da fala. Portanto, essa tecnologia vem de encontro ao desejo de muitos usuários, permitindo, até mesmo, aumentar o número de pessoas que utilizam os sistemas informatizados, pois reduziriam significativamente seu alto grau de rejeição.

Outro aspecto relevante, neste conjunto de soluções tecnológicas, foi o meio físico pelo qual a comunicação seria concebida. Qual é o equipamento doméstico que existe na maioria dos lares, e que é de tão fácil manuseio, que até uma criança sabe utilizá-lo? E que também tem uma infra-estrutura já instalada, capaz de ter tal abrangência que atenda o maior número de pessoas?

O elemento que preenche esses pré-requisitos é o aparelho telefônico, pois oferece a capilaridade necessária, graças a sua grande rede de telecomunicação, e desmistifica o acesso aos serviços disponíveis na Internet, dada a familiaridade das pessoas com esse tipo de equipamento. Com esse conjunto de soluções tecnológicas, a inclusão digital está acontecendo agora. Os portais de voz tornam isso possível, permitindo ao usuário executar todas as suas tarefas em apenas um único local. Os vários serviços pertinentes ao negócio dessa organização estarão centralizados e disponíveis para serem acessados através de comandos de voz, em um processo amigável e natural.

A perspectiva de unir as vantagens dessa nova tecnologia de reconhecimento de voz com os serviços oferecidos pelos governos aos seus cidadãos é o principal agente motivador desse trabalho. O grande número de pessoas beneficiadas por este tipo de solução, torna o portal de voz, com foco em governo eletrônico, um excelente objeto de estudo.

A quantidade desses serviços e a qualidade no atendimento dos mesmos serão obviamente melhoradas. Por exemplo, em setores como o de transporte, educação, saúde, habitação ou mesmo o de tarifas públicas, que atendem à maioria da população, serão beneficiados com a democratização na forma de acesso das informações. As mudanças refletirão na qualidade do atendimento, agilizando e diminuindo os transtornos causados à população, como a demora no atendimento e as filas intermináveis dentro das repartições públicas.



## 1.2 Objetivo

O enorme potencial existente nesse tipo de solução é indiscutível. O fato de trazer a "WEB" a todos os telefones foi impressionante, pois é a aliança da maior rede de dados mundial, a Internet, com o mais numeroso dispositivo de comunicação existente na Terra, o telefone. Entretanto, é de fundamental importância, que haja um crescimento na quantidade de aplicações existentes que utilizam a voz como forma de interface.

Os portais de voz preenchem vários requisitos necessários para a popularização do uso dessa solução, pois concentram em um único local, todos os serviços disponíveis por uma organização. A união desse tipo de portal, com a necessidade dos governos de prestarem serviços de melhor qualidade à população, serve como um excelente parâmetro, de como o bom uso dessa tecnologia pode ser bastante útil.

O principal objetivo deste estudo é propor a modelagem da interação do cidadão com o portal de voz como uma ferramenta essencial para facilitar o projeto da interação em portais de voz e alcançar os objetivos de usabilidade. Acredita-se que a modelagem resulte em um corpo de soluções mais consistentes do que seria obtido com o processo tradicional de projeto.

Este trabalho apresenta como estudo de caso a modelagem de um acesso típico do IPVA, Imposto sobre a Propriedade de Veículos Automotores da Secretaria dos Negócios da Fazenda do Estado de São Paulo do Governo do Estado de São Paulo e mostra como esta modelagem consegue antecipar problemas de usabilidade.

O modelo do serviço proposto insere-se numa perspectiva mais ampla, de um futuro portal de voz governamental, à imagem do que já existe no Governo do Estado de São Paulo no que se refere a portais de governo eletrônico para o cidadão.

Ainda, discutem-se a forma de autenticação e a identificação do usuário no sistema, que são notoriamente pontos chave para que se possa obter um processo seguro de acesso às informações.

A autenticação biométrica através da voz possui algumas vantagens intrínsecas, quando comparada a outras técnicas biométricas. Pode-se salientar a facilidade de captura da medida

biométrica (voz) utilizando "hardware" de baixo custo, a possibilidade de aquisição do sinal sem desconforto e sem a necessidade de contato físico do usuário (método pouco intrusivo), e a ausência de treinamento prévio para utilização do sistema. Esse método é o que será utilizado no portal de voz proposto, onde o processo de autenticação e identificação do usuário deverá ser solicitado apenas uma vez.

### 1.3 Contribuições do trabalho

Ao concluir satisfatoriamente o estudo em questão, um dos resultados esperados é a definição de um protótipo de um portal de voz para um sítio de governo eletrônico com foco no IPVA, no qual, serão oferecidos todos os serviços que o poder público normalmente disponibiliza para a população. Servirá também, como base para outros projetos, com características semelhantes, auxiliando os arquitetos de "software" na tomada de decisão.

Outra contribuição esperada, é que por esta tecnologia de reconhecimento de voz ser algo relativamente novo, no que se refere à implementação em aplicações comerciais, o estudo em questão ajudará na familiarização com os termos utilizados, as linguagens e padrões envolvidas no processo de desenvolvimento, a importância do uso de uma modelagem de interação, em especial a MoLIC, e por fim, permitirá conhecer algumas plataformas de desenvolvimento desse tipo de solução.

Portanto, a contribuição deste trabalho é, ao modelar um acesso típico no IPVA, fornecer subsídios para o desenvolvimento de projetos similares, que contemplem a fala como forma de interação entre o usuário e o sistema computacional. O conjunto desses procedimentos, permitirá conceber interfaces audíveis que contemplem a usabilidade, de tal forma que privilegie o seu uso, propiciando um acesso mais seguro e eficiente à informação desejada. Como consequência desse processo, haverá uma democratização dos serviços prestados por essas interfaces, beneficiando todos os usuários em potencial, como um grupo diferenciado composto por portadores de necessidades especiais, ou não.

### 1.4 Metodologia

A pesquisa relatada neste trabalho foi baseada em dois procedimentos, a partir dos quais foram

obtidos subsídios concretos para a execução do estudo em questão, a saber:

O primeiro passo foi uma pesquisa bibliográfica sobre os principais conceitos e sobre o estado da arte na área de interfaces audíveis e da tecnologia de reconhecimento de voz. As fontes de consulta mais importantes foram sítios sobre IHC e artigos de especialistas.

O segundo passo foi a análise dos diversos sítios de governo eletrônico existentes, com o objetivo de avaliar o grau de dificuldade da adaptação das interfaces visuais para uma interface audível.

A fusão dessas duas linhas de pesquisa contribuiu para a obtenção de um conjunto de informações concretas sobre os temas abordados e forneceu dados suficientes para as principais conclusões obtidas ao longo do trabalho. A incorporação da MoLIC para modelar o IPVA, com os seus conceitos de IHC permitiu identificar e enriquecer os principais processos interativos, observados durante a execução das tarefas propostas no portal de voz.

## 1.5 Organização do trabalho

Este trabalho está estruturado da seguinte forma:

A Seção 1 é esta Introdução.

A Seção 2 discorre sobre as Interfaces Audíveis, apresentando conceitos como o do som, das interfaces de fala e de não fala e as melhores práticas para o desenvolvimento de sistemas audíveis.

A Seção 3 fornece uma visão geral das ferramentas para modelar sistemas interativos e apresenta detalhes dos principais conceitos da MoLIC.

A Seção 4 aborda alguns aspectos comparativos, entre o portal visual e o portal de voz proposto. Apresenta, ainda, as linguagens e os padrões de desenvolvimento existentes para esse tipo de solução e um panorama geral das principais plataformas de desenvolvimento dessa tecnologia no país.

A seção 5 implementa a modelagem do IPVA para o portal de voz proposto e faz algumas

considerações sobre a contribuição da MoLIC na construção de aplicações audíveis.

A seção 6 apresenta dados conclusivos do estudo e as contribuições para trabalhos com essa mesma linha de pesquisa.

## 2. INTERFACES AUDÍVEIS

O fato das interfaces visuais ou gráficas serem muito mais utilizadas nos sistemas computacionais atuais fez com que os processos evolutivos das outras interfaces fossem mais lentos do que o desejado. O motivo é muito simples, pois a grande maioria dos usuários é altamente sensível aos estímulos visuais. Entretanto, foram surgindo diversas necessidades, com tal particularidade, que em alguns casos específicos, outros tipos de interface, demonstraram ser muito mais eficientes do que as visuais.

Dentro desse grupo especial existem as interfaces audíveis que podem ser definidas como sendo todo e qualquer processo de interação entre o usuário e o computador, através da utilização do som, sendo ele para a entrada ou saída de informação.

Ultimamente, alguns fatores têm estimulado o seu uso. Entre eles, estão as iniciativas sociais como a inclusão digital, que tem a grande meta de facilitar o acesso das pessoas as informações disponíveis nos sistemas computacionais. O que torna o uso da voz, como forma de comunicação, entre o ser humano e o computador, algo extremamente natural. O outro fator está diretamente ligado às características particulares de alguns aplicativos, por exemplo, situações em que o sistema computacional demanda uma grande sobrecarga visual, ou quando é necessário a tomada de decisão imediata, por parte do operador do sistema. Nesses casos, a utilização do som como forma de interação é bastante útil, pois fornece ao usuário todos os subsídios para que ele possa executar as suas tarefas de forma satisfatória.

Existem também, os casos em que a informação é apresentada em dispositivos com exibições visuais pequenas, como telefones móveis, ou "*personal digital assistants*" (PDAs). Nessas situações, as interfaces audíveis tornam-se fundamentais, proporcionando ao usuário uma forma mais fácil de obter a informação desejada. E por fim, o tema desse estudo, os portais de voz.

O enorme crescimento na utilização desses sistemas que contemplam a voz como meio de comunicação, reforçam o cuidado que os profissionais devem ter ao construir esse tipo de aplicativo. Para esclarecer e ressaltar essas afirmações, serão apresentados alguns dados que foram obtidos a respeito das características físicas e cognitivas dos seres humanos.

Segundo Bly (1982), experiências demonstraram que as respostas aos estímulos audíveis são mais rápidas do que os visuais. Sendo assim, existe um grande universo de aplicações a serem exploradas, como sistemas de alta criticidade que exigem tomada de decisão imediata e sistemas de tempo real. Nesta mesma linha de produtos existem os sistemas de respostas audíveis e sistemas por comando de voz.

## 2.1 Conceitos Gerais

Nesta seção, serão abordados temas como o processo de propagação do som, as principais características da audição humana, o som como forma de interação e as heurísticas para as interfaces audíveis.

### 2.1.1 Som

O texto a seguir tem o objetivo de facilitar o entendimento e a compreensão do leitor, aos termos e aos conceitos que são comumente utilizados ao longo do trabalho. Essa familiarização proposta, passa por uma breve recapitulação dos conceitos do som, até alguns detalhes do nosso sistema auditivo.

O som citado por Fonseca; Santos e Ferreira (2002) é uma energia vibratória (onda) produzida pelo atrito ou pela vibração de algum elemento físico com as moléculas de um elemento elástico qualquer. O vácuo perfeito não é um elemento elástico, já que significa a ausência de moléculas, sendo assim, impede a condução do som.

O elemento elástico condutor mais conhecido é o ar atmosférico, que conduz os sons que normalmente ouvimos, mas outros meios também servem para conduzir o som, como os sólidos, os líquidos e os demais elementos gasosos.

Um determinado som ao ser produzido, será conduzido pelas moléculas dos elementos elásticos que estão em seu caminho, normalmente o ar atmosférico. Essa energia vibratória (onda) é, então, conduzida por esses meios em forma de uma onda mecânica, até que toda a energia inicial seja dissipada em calor ou transformada em outra forma de energia.

Em cada elemento elástico o som terá uma velocidade diferente. Assim, o som no ar terá uma velocidade maior que o da água do mar, que será diferente de um outro meio qualquer.

As características do som que definem as suas qualidades são a altura, a intensidade e o timbre.

#### 2.1.1.1 Altura ou Tom

É o parâmetro que diferencia o grave do agudo; A altura, então, só depende da frequência sonora, que é o número de ciclos da onda sonora em 1 segundo. Quanto mais grave o som, menor a sua frequência. Quanto mais agudo maior a sua frequência. Ela pode ser medida por um frequencímetro e sua unidade de medida é o Hertz, cujo símbolo é Hz.

#### 2.1.1.2 Intensidade

É o parâmetro que diferencia os sons fracos dos sons fortes; É o fluxo de energia sonora por unidade de área. Como o processo de captação aural do ouvido humano não é linear, para essa medida utiliza-se uma escala logarítmica em unidades de decibéis (dB). Sons acima de 85 dBs podem causar danos permanentes ao ouvido, se houver exposição prolongada.

#### 2.1.1.3 Timbre

É o parâmetro que diferencia sons da mesma altura e de mesma intensidade emitidos por fontes diferentes. A mesma nota ré assume timbres diferentes se for emitida por um piano ou por um violão, já que cada nota produzida por um instrumento vem acompanhada de notas harmônicas diferentes. Assim o ré do piano vem acompanhado de harmônicos diferentes do ré do violão, permitindo a sua diferenciação pelo timbre.

### 2.1.2 Audição

A estrutura do ouvido, tanto pela sua complexidade, como pelo seu resultado operacional, é uma obra de extrema perfeição, e, por ser composto por um grupo de elementos altamente precisos, é considerado um aparelho bastante evoluído em termos de tecnologia biológica.

Ele possui três partes principais, ouvido externo, médio e interno, tendo cada um deles seu papel fundamental na identificação dos sons.

O pavilhão, que chamamos de orelha, possui curvas especialmente projetadas para fazer com que as ondas sonoras sejam dirigidas ao canal auditivo, levando-as em direção ao tímpano, que é a entrada do ouvido médio. Na maioria dos animais, por necessidade de sobrevivência, esse pavilhão é móvel, podendo ser alargado e orientado conforme a necessidade. No homem e na maioria dos animais considerados superiores essa função está prejudicada devido a sua rigidez. Esse pavilhão dirige as ondas para um ducto chamado de ouvido externo, que é fechado, em seu final, por uma membrana chamada de tímpano. O tímpano vibra com essas ondas sonoras, transmitindo essas vibrações a três minúsculos ossos ligados a ele, já no ouvido médio, que são o martelo, a bigorna e o estribo, amplificando-as e as transmitindo para o líquido existente no interior do ouvido interno.

No ouvido interno, que é constituído por um osso em forma de caracol chamado de cóclea, as ondas sonoras amplificadas pelos ossículos (martelo, bigorna e estribo), entram pela janela oval no canal vestibular até o final do canal, voltando pelo canal timpânico até sair novamente, desta vez pela janela redonda.

Durante todo o trajeto as ondas sonoras estão mergulhadas em um fluido aquoso chamado de perilinfa.

Esse fluido transmite as vibrações das ondas sonoras para o ducto coclear, que é o terceiro canal da cóclea, onde estão localizadas as células sensoriais ou células ciliadas (parecidas com pequenos pelos) que, quando agitadas pelas vibrações sonoras, produzem estímulos que transformam esse movimento em impulsos elétricos (energia eletro-química), que serão posteriormente decodificados pelo cérebro.

São milhares dessas células dispostas em todo o canal interno da cóclea e, cada uma é especialmente projetada para ser sensibilizada por uma vibração de uma só frequência, começando com aquelas que vibram com ondas de 30 vibrações por segundo e terminando com as ondas que alcançam 20.000 vibrações por segundo.

Por esse motivo é que os seres humanos não são capazes de ouvir todas as frequências sonoras, mas apenas aquelas que estão nessa faixa, que vai dos 30 Hz até os 20.000 Hz. Essa



energia eletro-química em forma de pulsos elétricos viaja até o córtex auditivo, no cérebro, por meio do nervo coclear. Só nesse momento é que começa a ocorrer o processo de entendimento sonoro, ou seja, só a partir daí é que ouvimos.

Portanto, a energia das ondas sonoras começa por ser de origem mecânica, como onda de pressão que se propaga primeiro no ar e depois num fluido, acabando na forma de impulsos elétricos que o cérebro utiliza para construir a imagem sonora correspondente.

### 2.1.3 O som como forma de interface

Ao combinar as informações fornecidas por dois dos sentidos humanos como o da visão e o da audição, as interfaces humano-computador, passaram a ser uma ferramenta altamente poderosa de interação. Na nossa vida cotidiana ambos os sentidos se complementam, provendo ao ser humano informações sobre o mundo exterior. A visão fornece uma informação mais detalhada sobre um foco concentrado, nos limitando a uma pequena área de atuação. Já a audição fornece um tipo de informação mais abrangente, nos alertando para as coisas que fogem da nossa visão periférica. A união desses dois sentidos nos permite obter todas as informações necessárias de um determinado ambiente.

A utilização desses dois sentidos na construção de sistemas computacionais pode proporcionar aos projetistas de interfaces uma excelente ferramenta de interação humano-computador. Segundo Blattner; Sumikawa e Greenberg (1989) existem várias vantagens ao se unir os sentidos da visão e da audição, quando na concepção de sistemas de multimídia e multimodal.

Ambos os sistemas estão diretamente ligados ao tema em questão. Os de multimídia são um tipo de interface que utilizam a imagem em conjunto com efeitos sonoros no desenvolvimento das aplicações, ou seja, provê ao usuário informações sobre eventos importantes e o auxilia a notar coisas que sob outras condições, poderiam passar por ele de forma despercebida. Os sistemas multimodais são um tipo de interface que fornecem informações integradas para ambos os sentidos, de forma mais eficiente. O foco principal é a utilização de alarmes e advertências, com o objetivo de monitorar e obter informações sobre o estado das atividades do sistema.

O mundo cinematográfico é um excelente exemplo da evolução dos processos de comunicação interativa ao longo dos anos. No início os filmes produzidos eram mudos e tentavam suprir esta carência na comunicação, através da ação e mímica dos seus personagens. Em uma fase posterior foram acrescentados alguns sons e temas musicais ao fundo do filme, proporcionando ao público sensações de tristeza, alegria, medo e outros. Atualmente não é possível conceber um filme sem contar com alguns desses itens de comunicação audível, como músicas, sons e efeitos especiais, permitindo ao público ter uma verdadeira imersão nas sensações sonoras tridimensionais.

As Interfaces Humano-Computador passaram por esse mesmo processo evolutivo, no qual inicialmente eram pobres tanto nos aspectos visuais como nos audíveis e ao longo do tempo foram sendo enriquecidas através da utilização de técnicas de multimídia e de multimodal.

É importante notar que da mesma forma que os sons nos auxiliam no mundo cotidiano, eles também podem ser muito úteis nos sistemas computacionais, porém cabe observar que esta sugestão deve ser vista com alguma reserva. Segundo Barrass(1998) quando esses dispositivos sonoros são mal implementados poderiam tornar o computador uma ferramenta bastante ruidosa e incômoda. Esta preocupação realça a necessidade de projetar interfaces que produzam sons úteis, ao invés de novidades divertidas ou aplicativos barulhentos que provoquem a dispersão do usuário da tarefa principal e perturbem o ambiente de trabalho.

## 2.2 Utilização de Interfaces Audíveis

Conforme Barrass (1998) a tomada de decisão pelos projetistas, isto é, em que momento optar por uma interface audível, em detrimento de qualquer outro tipo de interface, está diretamente relacionada a um conjunto de circunstâncias específicas que devem ser observadas. A seguir será apresentada uma lista que pretende nortear os profissionais no momento dessa definição, a saber:

- Quando a interface utilizada não é visual.
- Quando o conteúdo da mensagem é utilizado no momento em que ocorre o evento.
- Quando a origem do sinal é um som.
- Quando a mensagem é simples e curta.

- Quando é enviado um sinal de advertência ou uma determinada mensagem pede uma ação imediata.
- Quando existe a necessidade de apresentar informações contínuas e variáveis de algum tipo, como aeronave, comunicação por rádio, ou informação de rota de vôo.
- Quando a ação relativa à mensagem não pode ser executada mais tarde.
- Quando o sistema visual é sobrecarregado.
- Quando são empregados unicamente canais de fala na interação.
- Quando a iluminação é insuficiente e limita a visão.
- Quando o equipamento é uma interface móvel.

O conjunto de sugestões acima relacionado, certamente não abrange todas as possibilidades existentes, contudo proporcionarão aos desenvolvedores uma noção de qual rumo deve ser seguido quando houver a necessidade da tomada de decisão, entre o uso ou não de uma interface audível em um sistema computacional.

### 2.3 Heurísticas para as interfaces audíveis e Portais de Voz

As aplicações audíveis desenvolvidas na área militar para controle de aeronaves e orientação dos operadores estão bastante adiantadas. Documentos como o *Design Criteria Standard Human Engineering* do Department of Defense - EUA (1999), apresentam alguns de seus métodos, procedimentos e padrões, que servirão para nortear e auxiliar os desenvolvedores durante o processo de construção de um sistema audível. Dada a extensão do documento e um grande número de opções ser de pouca relevância para as aplicações comerciais, a abordagem será baseada no conjunto de itens descritos a seguir:

\* Sinais de advertência e alarmes:

Os sinais devem ser usados para advertir o usuário do perigo, alertá-lo de mudanças críticas no estado do sistema ou do equipamento e avisá-lo da necessidade de tomada de ação imediata. Os sinais são extremamente úteis para detectar situações que a observação normal não forneceria com a mesma eficiência. Podem ser sub-divididos quanto a sua:

Natureza: Deve ser compatível com o tipo de alerta a ser fornecido ao usuário, normalmente

seguem padrões pré-determinados;

Frequência: Deve estar na escala de 200 a 5.000 Hz, preferencialmente de 500 a 3.000 Hz e não deve ultrapassar 115 DB;

Intensidade: Deve ser compatível com o ambiente acústico;

Criticidade: Deve ser repetida a cada 0,5 segundos até ser corrigida ou ser desativada.

\* Mensagens de advertência por voz:

As mensagens são utilizadas para fornecer ao usuário informações mais detalhadas e específicas, em casos que os sinais não seriam suficientemente explicativos. Podem ser subdivididos quanto a sua:

Natureza: Devem ser objetivas, claras e padronizadas, deixando bem definidas quais as ações Devem ser tomadas ;

Intensidade: Deve ser no máximo 115 DCB;

Tipo de voz: deve ser clara e madura;

Estilo: Deve ser formal e impessoal;

Uso: Deve somente ser utilizada no caso de absoluta necessidade, com o objetivo de esclarecer a ação a ser tomada;

Forma da mensagem: Deve ser elucidativa, descritiva e concisa respeitando esta ordem. Também para facilitar a compreensão à mensagem deve ter no máximo quatro sílabas, a menos que o seu resultado seja inconsistente;

Criticidade: Deve ser repetida a cada 3 segundos até ser corrigida ou ser desativada.

Cabe notar que em ambos os casos acima mencionados o som da interface deve ultrapassar 15 DCB o ruído do ambiente, para que possa ser ouvida pelo usuário; Contudo, não deve ultrapassar 25 DCB nessas mesmas condições ambientais, pois se tornaria muito alto causando perturbação e desconforto ao usuário.

Os portais de voz também pertencem ao grupo de aplicações audíveis e as heurísticas acima referenciadas, tem um grande peso no seu desenvolvimento. Entretanto, existe um conjunto de

procedimentos mais específicos para esse tipo de aplicação. Essas heurísticas foram agrupadas, com o objetivo de auxiliar o entendimento dessas melhores práticas e facilitar as futuras referências. A saber:

a) Balanceamento na árvore de opções

Devem ser o suficientemente verticalizadas (lista de opções na vertical) para aproveitar ao máximo essa interação, sem provocar o esquecimento das opções já dadas ou a dispersão do usuário. O mesmo ocorre no caso da horizontalização (lista de opções na horizontal). O balanceamento desses dois itens, faz com que a interação não se torne demasiadamente longa ou cansativa, para se atingir o objetivo final . (Gruenstein, 2002).

b) Fraseologia

Deve ser clara e objetiva. Sem que seja muito curta, pois se torna de difícil compreensão, e também, não pode ser muito longa, pois provoca o cansaço e a dispersão. Nesta heurística, também deve ser considerada a altura, o timbre e a clareza da voz que está sendo utilizada no processo de comunicação com o usuário , citado por Burkhardt; Stegmann e Van Ballegooy (200-?).

c) Linguagem

Deve ser muito próxima da linguagem natural do usuário. O objetivo dessa heurística é o de evitar siglas, abreviações e termos técnicos que provocam dificuldade no entendimento das opções disponíveis, citado por Pakucs (19--?).

d) Suporte ao Usuário

Deve possibilitar ao usuário corrigir falhas, sair do sistema quando desejar e prover ajuda a qualquer momento durante o processo de interação, citado por Hone e Graham(2001).

e) Flexibilidade

Deve permitir ao usuário retornar às opções anteriores, auxiliar o usuário quando não houver entrada no diálogo (noinput) e prover acessos rápidos aos serviços disponíveis (atalhos), citado por Kondratova (2004).

f) Confiabilidade

Deve fornecer ao usuário, a segurança necessária no que se refere a privacidade e integridade das informações. Também deve propiciar acessos rápidos (pequeno intervalo de tempo entre as interações), possibilitar uma alta taxa de reconhecimento da fala humana e por fim, não perder o contato com o portal de voz (estabilidade na ligação telefônica) , citado por Branco et al. (2006).

#### 2.4 Interfaces audíveis de fala

Segundo Robinson (1996) em "*Speech Analysis*", as interfaces audíveis de fala ganharam um grande impulso através dos crescentes avanços tecnológicos que fizeram com que o reconhecimento da fala se tornasse um campo de estudo fascinante e ao mesmo tempo desafiador, uma vez que é bastante grande a gama de aplicações, onde o uso da voz tem o papel de agilizar e facilitar a realização de tarefas cotidianas, buscando extrair da fala as informações relevantes para a realização das mesmas.

A interação com este tipo de interface é muito simples e eficaz, pois é o meio natural que o ser humano utiliza para expressar seus pensamentos, de tal forma que humaniza mais o relacionamento entre o homem e o computador. Na atualidade existe uma tendência mundial em produzir equipamentos com interfaces cada vez mais amigáveis, o que na realidade demonstra uma intenção de inverter a cadeia de prioridades da indústria, que demorou tempo demais para descobrir que não é o homem que deve se adaptar ao funcionamento das máquinas, e sim o contrário. Uma prova desta mudança de paradigma está na afirmação dos pesquisadores responsáveis pelo projeto Oxygen do Instituto de Tecnologia de Massachusetts, o MIT que segundo a revista Inovação Tecnológica (2002), afirmaram: "Nos primeiros 40 anos da ciência da computação, estivemos preocupados em adaptar a tecnologia àquilo que as máquinas querem". Atualmente o projeto Oxygen, que foi criado por uma associação do laboratório de ciência da computação e o laboratório de inteligência artificial, ambos do MIT, promete apresentar em 5 anos, equipamentos e programas tão fáceis de acessar quanto "o ar que se respira".

O artigo "A voz do dono e o dono da voz" de MENCONI (2004?) publicado em uma revista de grande circulação no Brasil, destacou que pesquisadores da Unicamp, desde 1994 estão estudando as particularidades da língua portuguesa. Hoje a universidade tem um dos mais

elogiados sistemas de reconhecimento e sintetização da fala humana do País. São 239 mil palavras cadastradas num banco de dados único, incluindo aí as conjugações verbais e mil exceções gramaticais. Cabe ressaltar que a criação de um sistema de reconhecimento de voz não é uma tarefa simples. Não adianta transferir o conteúdo do dicionário para o computador. É necessário repetir as mesmas palavras com diferentes sotaques, dialetos e pronúncias regionais. Para ser produtivo, um sistema de reconhecimento de voz deve acertar pelo menos 95% das palavras ditadas em fala contínua.

#### 2.4.1 Sistema de Reconhecimento de Voz

O reconhecimento de voz possibilita uma manipulação mais fácil e prática de equipamentos providos com a capacidade de compreender a fala humana. Graças aos avanços tecnológicos na área de "*software*" e "*hardware*" esta tarefa torna-se cada vez mais possível de ser realizada.

As técnicas de processamento de sinais permitem a extração de características que realmente mereçam destaque, pois atuam no sentido de fornecer não somente a informação de interesse ao processamento de determinada amostra de som, como também ocasionar uma redução considerável na quantidade de informações a serem processadas. Tais informações serão responsáveis pela produção de padrões a serem identificados numa comparação entre determinada referência registrada e a apresentação de uma nova amostra para teste. Então, o papel do sistema será validar ou não determinada amostra, dependendo do tipo e funcionalidade do sistema de reconhecimento em questão.

Conforme Proakis; Deller e Hassen (1993) existem vários métodos para comparação entre padrões, os quais serão destacados os de maior relevância como o Alinhamento Temporal Dinâmico (*Dynamic Time Warping*, DTW), os Modelos Escondidos de Markov (Hidden Markov Models, HMM) e as Redes Neurais Artificiais (Artificial Neural Networks, ANN). Esses sistemas podem ser utilizados individualmente ou em conjunto no processo de reconhecimento da fala.

Segundo Valiati (2000), os sistemas de reconhecimento de voz podem ser divididos em relação a forma de como são tratadas as palavras no momento da identificação, como sendo palavras isoladas, concatenadas ou contínuas, onde deve ser considerado um intervalo de

silêncio no tempo entre a pronúncia de cada palavra a fim de diferenciar esta classificação. Também são consideradas as pronúncias dependentes de determinado locutor, assim como as que não fazem distinção entre locutores, importando somente o que foi dito e não quem o disse.

Dentre as aplicações que utilizam o reconhecimento de voz, existem duas que merecem algum destaque, devido ao seu uso e importância que atualmente representam para as interfaces audíveis de fala.

A primeira aplicação são os portais de voz, que por definição são aplicações que acessam dados no mundo da Internet de maneira transparente, como ouvir mensagens do seu correio eletrônico pelo telefone. Permitem também a navegação através de palavras chave ou de teclas de atalho, e em uma única ligação, o sistema deve permitir ao usuário acessar várias informações e dados.

Outra técnica que tem demonstrado ser de grande interesse, para as aplicações de reconhecimento de voz é a de identificação pela fala. Ela normalmente está dividida em duas etapas distintas que são o treinamento e o reconhecimento. No treinamento, são gravadas frases foneticamente balanceadas ditas pelo usuário. A voz é digitalizada para que sejam extraídas suas características, como a frequência fundamental, que dá o tom da voz, e as frequências de ressonância do aparelho fonador. Em seguida, essas características são transformadas em conjuntos de matrizes e números, que passam por um software que simula matematicamente o funcionamento do cérebro. Surge desse processo uma nova matriz, definitiva, que vai identificar a voz de cada usuário.

Na etapa do reconhecimento o usuário se identifica e tem a voz analisada. Se as características extraídas forem iguais e a matriz a mesma, o sistema reconhece a voz do usuário.

O processo é bastante seguro, pois a voz humana é como a impressão digital, ou seja, quase impossível de se falsificar. A singularidade de cada voz se deve à estrutura do aparelho fonador, das cordas vocais e do formato dos lábios. Como as frequências de ressonância resultantes são diferentes para cada indivíduo, torna-se impossível reproduzi-las.

#### 2.4.2 Unidade de Resposta Audível



A URA (Unidade de Resposta Audível) ou telefonia computadorizada (CTI -*Computer Telephony Integration*) é uma plataforma de produtos de processamento de chamadas que recebe voz ou dígitos, e possibilita uma série de aplicações via voz e/ou fax, tais como saldo bancário, informações ao público, retirada de pedidos, status de mercadorias, entre outros. Esse sistema oferece serviços e facilidades de forma automática, permitindo ao usuário navegar por menus interativos através do telefone e obter uma resposta rápida para uma consulta a um banco de dados, sem a interferência humana, o que traz grande economia em pessoal e presta um ágil e eficiente atendimento ao usuário.

As URAs fornecem serviços de interação com o usuário através da entrada de dados pelo teclado telefônico ou por sistemas de reconhecimento de voz. Inicialmente, esse tipo de produto apenas suportava a entrada de informações através do teclado telefônico, porém com o avanço da tecnologia do reconhecimento da fala, os usuários passaram a contar com este agente facilitador no uso, ampliando de forma significativa o acesso a esse tipo de serviço.

A grande vantagem da unidade de resposta audível é que o equipamento utilizado para fazer a comunicação entre o usuário e o sistema é um simples telefone, que até uma criança sabe operar, e uma linha telefônica, o que se tornou bastante acessível nos últimos anos, dada a evolução da estrutura de telecomunicações no Brasil. Outro ponto favorável é a forma de interação, que é através de mensagens ou frases curtas não havendo a necessidade de treinamento prévio para sua utilização. Se a plataforma for a que recebe voz, o usuário conta com a facilidade do reconhecimento inteligente e natural das palavras, e é orientado a repetir a fala caso não seja compreendido.

Os principais recursos fornecidos pelas URAs são:

- . Atende ligações telefônicas 24 horas por dia e 7 dias por semana;
- . Administração da fila de espera;
- . Pode ser conectado a uma ampla variedade de bancos de dados;
- . Pode conter recursos de fax e suportar interfaces analógicas e digitais;
- . Reproduz vozes digitalizadas, como gravação de mensagens institucionais ou árvores de menu;
- . Recebe dígitos DTMF e decádicos (pulso);
- . Recebe e interpreta voz;

- . Responde via voz ou fax, acessando um banco de dados ou um *host* (servidor);
- . Transfere ligações;
- . Disca números telefônicos.

Cabe observar que o conceito das unidades de respostas audíveis deve ser revisto. Situações onde apenas seria fornecida uma lista contendo um conjunto estático de opções de serviços, já não é uma verdade. A interação através de comandos de voz, em alguns casos substituiu totalmente o uso da árvore de menus, aumentando de forma imensa a variedade e diversidade dos serviços oferecidos aos usuários. Somente para exemplificar, existe o serviço de auxílio à lista (102) oferecido pela Telefônica e o serviço de informações sobre a conta corrente com o uso do telefone, oferecido pelo banco ITAÚ. Em ambos os casos, o acesso a informação desejada, é obtida através da navegação entre as várias opções do aplicativo, por comandos de voz.

A prática tem demonstrado, que atualmente as empresas estão fornecendo sistemas híbridos, que contemplam tanto a entrada de comandos através de dígitos como por voz. Isto se explica devido a duas circunstâncias, a primeira é que o reconhecimento da fala é uma nova tecnologia e como tal gera um alto custo financeiro na sua implantação. E a segunda é que os usuários ainda não se sentem confortáveis ao ditar suas senhas bancárias e outras informações sigilosas.

#### 2.4.3 Sistema de Reprodução da Fala

A tecnologia de reprodução da fala citado por Constantine(2003) apresentou algumas características especiais, que proporcionaram duas vertentes diferentes em relação ao seu uso e forma, que foram se adequando as necessidades individuais e específicas de cada uma delas.

O primeiro foi o sistema de sintetização de voz que está baseado em um processo de interpretação das palavras contidas em um texto e na sua conversão no momento da execução para sons sintetizados, que correspondem aos fonemas necessários para a verbalização. Esse tipo de sistema é muito utilizado em aplicações onde a verbalização necessita ocorrer de forma dinâmica, isto é, gerar um conjunto de palavras e frases fora de um contexto pré-definido. Os usuários portadores de necessidades especiais são um dos principais grupos beneficiados por essa solução tecnológica, pois os programas que permitem sua comunicação

com o computador utilizam esse tipo de sistema no seu desenvolvimento, para permitir o acesso desses usuários ao mundo da informática. Entre eles pode-se citar programas como o *Jaw*, *Slimware Windows Bridge*, e o *Virtual Vision*, entre outros que utilizam o conceito de sintetização dinâmica da fala humana.

O segundo foi o sistema de concatenação de fonemas, que está baseado na montagem de frases e palavras previamente gravadas dentro do sistema. A análise do texto a ser convertido ou a execução de um evento associado a um serviço, dispara uma pesquisa em um conjunto de fonemas já pré-definidos e armazenados em um banco de dados interno do sistema. A concatenação desses fonemas produz a verbalização desejada. Este tipo de aplicação é muito utilizado em conjunto com as URAs, onde permitem ao usuário obter respostas as suas solicitações em sistemas de saldo e extrato eletrônico, entre outros.

As perspectivas das interfaces audíveis de fala são as mais otimistas e apontam para uma larga gama de utilização em sistemas de segurança biométrica, portais de voz e interação verbal entre o ser humano e o computador. Geralmente, em uma plataforma de atendimento automatizada, são utilizadas três etapas durante o processo de comunicação conforme citado no artigo APLICAÇÕES com Reconhecimento de Voz , a saber:

a) Reconhecimento da fala ASR - (*Automatic Speech Recognition*), é o principal elemento da solução, sendo classificado em dois tipos:

\* Sistemas dependentes do usuário (*speaker dependent*), no qual reconhecem apenas a fala de determinado indivíduo (como os usados em agendas de telefones celulares);

\* Sistemas independentes do usuário (*speaker independent*), no qual reconhecem padrões aplicáveis a diversos indivíduos, sendo adequados para utilização no atendimento a inúmeras pessoas.

b) Autenticação de voz, que possibilita identificar quem é o usuário, sem a necessidade do uso de senhas. A voz é uma biometria humana e sua autenticação baseia-se na análise de padrões harmônicos e não simplesmente na comparação entre reproduções de uma mesma fala, sendo uma alternativa extremamente segura contra tentativas de fraude.

c) Conversão de texto para voz TTS - (*text-to-speech*), é o elemento da solução que permite a verbalização de textos. Isto viabiliza a comunicação através da fala humana, que permite retornar a informação solicitada pelo usuário.

\*Na sintetização, o algoritmo analisa o conjunto de palavras do texto e sintetiza os sons correspondentes aos fonemas necessários para a verbalização. Por exemplo, a possibilidade de ouvir um e-mail através do telefone;

\*Na concatenação, o algoritmo analisa o texto submetido e define uma seqüência de fonemas pré-gravados que, quando concatenados, reproduzem em som o texto em questão.

Quanto à conversão de textos para voz, cabe fazer algumas considerações sobre os dois métodos apresentados. Na concatenação, que é o método mais utilizado, são apresentadas algumas desvantagens, como por exemplo, a fala descontínua e a limitação na quantidade de textos a serem reproduzidos. Como vantagem, pode-se destacar o custo da solução, que apresenta aspectos tecnológicos mais simples e comumente encontrados no mercado. Já na sintetização, a vantagem é que a fala é contínua e não existe a limitação nos textos, pois a reprodução da fala é executada de forma dinâmica, o que proporciona maior flexibilidade para a solução. Como desvantagem, pode-se destacar o seu custo, que depende de alguns aspectos tecnológicos mais complexos.

## 2.5 Interfaces audíveis de não fala

As interfaces de não fala tem como principal forma de interação, sons artificiais de alertas e alarmes, que podem ser produzidos por diversas fontes como o som de instrumentos musicais sintetizados, ou também, sons naturais sintetizados como o rasgar de papéis. A variação da sua frequência, timbre e tonalidade geram tons que permitem criar espaços virtuais audíveis. Isto é muito útil, pois possibilita aos desenvolvedores optar por uma enorme gama de tons que podem ser associados a várias tarefas e eventos dentro de um sistema computacional.

Este tipo de interface oferece algumas vantagens interessantes como possibilitar ao usuário continuar executando sua tarefa principal enquanto paralelamente interage com o sistema através de sinais sonoros. Para exemplificar o exposto, serão descritas duas situações de uso, a saber:

\* O usuário está executando uma tarefa demorada, como a instalação de um novo produto ou o "download" de um arquivo da Internet. Em condições normais ele necessita ficar monitorando a barra de progressão sendo obrigado a desviar sua atenção, de vez em quando, para verificar se a tarefa foi executada. Caso houvesse uma barra de progresso sonora, que indicasse através de vários tons diferentes sua evolução e o seu término, o usuário poderia em paralelo editar um documento ou executar qualquer outra tarefa.

\* O usuário ao utilizar uma interface móvel normalmente está em trânsito, isto significa que toda sua atenção deve estar focada na tarefa principal, que é a de andar ou a de fazer uma outra atividade qualquer. Sendo assim, o uso de interfaces audíveis de não fala como forma de interação entre o usuário e o equipamento móvel, faz com que não ocorra a total dispersão na execução dessa importante tarefa, permitindo ao usuário executar as duas atividades em paralelo.

### 2.5.1 Ícones Audíveis

Os ícones audíveis foram idealizados por Gaver (1986) e são definidos por ele como sendo efeitos sonoros dentro de um sistema computacional, que representam sons naturais e fazem parte do nosso dia a dia. Portanto, não necessitam ser aprendidos e podem ser utilizados para representar ações e objetos dentro de uma interface.

Os ícones audíveis estão baseados em uma analogia entre o nosso mundo cotidiano e o mundo do computador.

### 2.5.2 Earcons

Os *earcons* foram desenvolvidos por Blattner; Sumikawa E Greenberg(1989) e são definidos como sendo sons sintéticos e abstratos que utilizam tons em combinações estruturadas para criar mensagens audíveis. Portanto, necessitam ser aprendidos. Os *earcons* estão baseados em notas musicais e são mensagens auditivas não verbais, utilizadas em uma interface com o objetivo de prover informações para o usuário sobre algum objeto, operação ou interação em um sistema computacional.

### 2.5.3 Técnicas de Exibição

As técnicas de exibição audíveis foram classificadas por Blattner, Papp e Glinert (1994) como sintáticas, semânticas e léxicas.

#### 2.5.3.1 Técnica sintática

A ênfase da técnica sintática, está baseada na organização entre os sinais audíveis. O código Morse é um exemplo característico deste tipo de técnica, pois utiliza a variação na duração e no ritmo do som para enviar mensagens de texto codificadas.

Outro representante desse tipo de técnica são os *earcons* e baseiam-se em sons sintéticos, que permitem criar estruturas hierárquicas com diferentes tons e notas musicais. Eles são compostos de componentes que podem variar em ritmo, tom, timbre, métrica e dinâmica. Cada *earcon* tem um significado único e deve ser aprendido, por exemplo, um tom com a frequência de 400 Hz pode significar "arquivo" e o mesmo tom com a frequência de 700 Hz pode significar "exclusão". Como também pode ser criado um *earcon* composto, ao combinar os dois tons anteriormente descritos, significando "exclusão de arquivo".

A grande vantagem da técnica sintática é que os *earcons* podem ser construídos com muita facilidade e permitem ser reproduzidos em qualquer equipamento que possua uma placa de som para ouvir músicas ou que contenha ferramentas de manipulação audível. Eles não necessitam ter nenhuma associação natural com os objetos que irão representar e estes signos podem ser criados pelo som de instrumentos musicais como piano, violão, saxofone, flauta e outros. A desvantagem é que como os *earcons* devem ser aprendidos, sua eficiência está diretamente relacionada com a capacidade que o usuário tem de associar a execução da tarefa desejada com o som que ela representa. Outra coisa a ser observada é que os *earcons* quando são compostos de dois ou mais tons, podem ser extremamente difíceis de ser aprendidos.

#### 2.5.3.2 Técnica semântica

A ênfase na técnica semântica, está baseada na representação do significado de cada som. Os ícones audíveis são os grandes representantes dessa técnica e podem ser definidos como sendo signos que permitem ao usuário fazer uma associação das tarefas a serem executadas com alguns sons do mundo cotidiano. Gaver (1994) sugere que os ícones audíveis por utilizarem modelos do mundo real são de fácil aprendizado, pois os seres humanos estão acostumados a

receber as informações através deste tipo de som.

A criação de um ícone audível dentro de um sistema computacional, começa com uma análise da interação entre um objeto da interface e sua eventual associação com um som do mundo real. Por exemplo, o arrastar de um arquivo de uma pasta para outra dentro de uma interface audível, seria representada pelo som real do arrastar de uma pasta, de um lugar para outro, sobre a superfície de uma mesa.

A vantagem do modo semântico é que os ícones audíveis são mais intuitivos do que os *earcons*, mas a desvantagem é que nem sempre é possível encontrar uma associação de um objeto da interface com um som do mundo real. Segundo Lucas(1994) situações empíricas demonstraram não haver diferenças significativas de aprendizado entre ícones audíveis e *earcons*. Portanto, cabe concluir que a afirmação mais coerente para o fato em questão é o de que alguns estudos sobre o aprendizado e identificação dos ícones audíveis têm demonstrado que a sua interpretação depende da experiência, da expectativa e do contexto dos seus usuários. Embora a técnica esteja baseada na teoria da percepção inata do ser humano e deveria assegurar uma fácil associação entre os objetos, resultados experimentais têm demonstrado diferenças individuais significativas.

#### 2.5.3.3 Técnica léxica

A ênfase da técnica léxica, está na forma pela qual o signo é conhecido. Fornece subsídios para evitar confusões e provê formas para distinguir sons e definir seus significados, permitindo ao usuário tomar as ações corretas quando necessário. Esta técnica é bastante prática e se preocupa com alguns parâmetros, taxas e níveis dos sons, quando utilizados para a emissão de alertas e mensagens dentro de uma interface audível.

A técnica léxica tem como principal objetivo, evitar que a comunicação seja prejudicada por fatores externos, como ruído ambiental, distância entre a interface e o usuário, forma da mensagem, frequência e intermitência do alerta. A quantidade dessas recomendações é muito grande e sua abrangência é extensa, porém apenas a título de ilustração, segue uma recomendação léxica, que orienta utilizar frequências abaixo de 1000 Hz quando houver uma grande distância entre a interface audível e o usuário.

#### 2.5.4 Avaliações Audíveis

Quanto as avaliações auditivas, pode-se afirmar que vários estudos utilizando *earcons* e ícones audíveis estão sendo feitos. O objetivo destas avaliações, em geral, tem sido obter interfaces audíveis que permitam ao usuário executar as tarefas de interação com o equipamento móvel, sem prejudicar a sua atividade principal, que normalmente é a de caminhar.

##### 2.5.4.1 Espaço audível em 2D

Experiências têm avaliado a eficiência no uso de interface de não fala simulando espaços virtuais em 2D, citado por Brewster; Lumsden e Bell et al. (2003) . O processo baseia-se no movimento dos dedos em uma tela de um PDA como um mecanismo para selecionar opções em um sistema computacional. Para tanto, foi utilizado um equipamento PDA preso ao cinto do usuário e fones de ouvido, onde a comunicação entre ambos é executada através de uma rede sem fios.

A navegação entre as opções do menu torna-se possível, através do deslizar dos dedos do usuário sobre a superfície da tela do PDA. Já a identificação do item desejado é obtido com o auxílio de ícones audíveis e *earcons*, que são transmitidos para o fone de ouvido. Após essa identificação, basta o usuário pressionar com o dedo, o local na tela, para selecionar o item escolhido.

As interfaces audíveis de não fala em 2D, demonstraram ser uma grande alternativa para a manipulação de equipamentos móveis, porém algumas dificuldades foram encontradas pelos usuários durante o seu uso. O fato de caminhar e ao mesmo tempo escolher a opção desejada através do deslizar do dedo pela tela, provocaram algum desconforto, gerando assim perdas de eficiência na comunicação entre o usuário e o sistema computacional.

##### 2.5.4.2 Espaço audível em 3D

Experiências têm avaliado a eficiência no uso de interface de não fala simulando espaços virtuais em 3D, citado por Brewster; Lumsden e Bell et al. (2003) . O processo baseia-se no movimento de cabeça como um mecanismo para selecionar opções em um sistema



computacional. Para tanto, foram utilizados rastreadores de movimento da cabeça e fones de ouvido. Os menus não visuais foram criados de forma circular divididos em quadrantes, onde através do auxílio de ícones audíveis e *earcons*, o sistema fornece subsídios para que o usuário navegue com relativa segurança entre as opções do menu.

Após a identificação do item desejado, que é obtido através do movimento da cabeça da esquerda para a direita e de cima para baixo, um outro movimento de cabeça curto e rápido faz a seleção da opção desejada, simulando o "*click*" do "mouse". Deve-se notar que todas essas funções foram executadas com os usuários caminhando e os resultados atingidos podem ser considerados satisfatórios.

Esta seção apresentou uma visão geral dos principais conceitos e o atual estágio das interfaces audíveis de fala e de não fala. Também abordou as melhores práticas a serem aplicadas, durante o processo de desenvolvimento de uma interface audível e as suas tendências futuras. A próxima seção esclarece o leitor sobre a importância das Linguagens de modelagem interativas e fornece os principais conceitos da MoLIC.

### 3. MODELAGEM PARA SISTEMAS INTERATIVOS

A concepção de um novo sistema segue várias fases durante o seu processo de desenvolvimento, já bastante abordadas na Engenharia de Software. Entre elas, existe a modelagem de interação que é considerada uma fase muito importante dentro dos conceitos de IHC. Normalmente, ela acontece durante a fase de elaboração, quando os requisitos são definidos.

A modelagem de interação se preocupa com o processo de comunicação entre o usuário e o sistema computacional. Para tal, utiliza todos os recursos disponíveis para facilitar esse diálogo e contemplar a usabilidade e a acessibilidade requeridas pela aplicação.

A MoLIC é uma linguagem que pode fornecer todos os subsídios necessários para se atingir esse objetivo. Os seus processos estão baseados nos princípios da Engenharia Semiótica, que privilegiam vários aspectos na comunicação entre o usuário e o aplicativo. No caso do portal de voz, a implementação de uma modelagem para sistemas interativos, beneficiará de forma significativa o seu uso.

#### 3.1 Semiótica

A semiótica, ciência que estuda os signos e a sua vida na sociedade, vem sendo aplicada a estudos de IHC e pode ser uma contribuição importante à evolução dos processos de desenvolvimento das interfaces audíveis. Ela teve origem há mais ou menos dois mil anos, conforme citado por Eco (1976) e muito tem sido escrito sobre sua aplicação a diversos campos do conhecimento humano.

De acordo com (SANTAELLA, 1996) "um signo é qualquer coisa que está no lugar de outra para alguém sob determinadas condições" e a sua contribuição em IHC é devido ao fato de ser uma disciplina que estuda a expressão e comunicação, permitindo considerar uma interface audível como um ambiente virtual e abstrair sensações de envolvimento, interação e exploração do mundo representado por uma interface. Sendo assim, a engenharia semiótica fornece conceitos que quando aplicados no desenvolvimento de sistemas proporcionam a forma mais avançada de interface com o usuário, que é a possibilidade de imersão.

Segundo Oliveira (2000) essa capacidade de imersão é obtida na engenharia semiótica ao se tratar uma interface como sendo um conjunto de entidades que se comunicam entre si, uma ou mais das quais sendo seres humanos. Os processos comunicativos entre essas entidades são fenômenos de natureza semiótica. Portanto, uma entidade pode ser identificada, dentro do mundo da interface por uma característica que a individualiza. Como por exemplo, uma entidade em uma interface audível pode ser identificada por um tipo de som específico ou por uma tonalidade de voz.

### 3.2 MoLIC – Modeling Language for Interaction as Conversation

Existem vários elementos que comprovam a crescente evolução da engenharia semiótica dentro dos processos de desenvolvimento de sistemas com base nos conceitos de IHC. Entre eles, destacam-se iniciativas como a de um grupo de profissionais do Departamento de Informática da PUC-RIO, que desenvolveram uma linguagem de modelagem denominada MoLIC, citado por Barbosa; Souza e Paula (2003). Segundo as autoras, MoLIC foi concebida com base nos princípios da engenharia semiótica em conjunto com os conceitos de IHC. Ela define uma interface como sendo uma metamensagem enviada do aplicativo para os usuários, e vice-versa. Essa mensagem é criada de tal forma que é capaz de trocar mensagens com os usuários, possibilitando a interação entre o usuário e o sistema. Na engenharia semiótica, uma interação é definida como sendo um processo de conversação entre duas ou mais entidades envolvidas nessa comunicação. MoLIC utiliza alguns métodos próprios que a auxiliam a criar seus modelos de interação e de comunicação.

A MoLIC representa um grande avanço no processo de modelagem dos novos sistemas computacionais e implementa vários conceitos de IHC. Sendo assim, é utilizada no estudo em questão, com o objetivo de enriquecer as etapas de construção do portal de voz com foco em governo eletrônico. Na seção 5, onde é apresentada uma modelagem do IPVA, são utilizados os cenários, o modelo de tarefas, o modelo de interação e todos os seus processos derivados recomendados pela linguagem. Estes processos são descritos a seguir, com maior detalhe, com base no trabalho de definição original da linguagem.

### 3.3 Cenários

Segundo Silva e Barbosa (2004) os cenários na MoLIC são descrições minuciosas de algumas

situações de uso do aplicativo. Eles podem simular situações de interação (cenas) reais ou não. Os cenários refletem ações executadas pelo usuário, pelos processos do aplicativo e pelas informações trocadas entre ambos. Entretanto, eles não devem conter detalhes dos objetos do diálogo, como botões, rótulos, caixa de textos, entre outros. Isso permite que o projetista não seja influenciado pela solução tecnológica da interface, sem antes passar pelas etapas de construção do modelo de tarefas e de interação.

Os cenários são importantes para que possamos identificar as principais funcionalidades e interações do aplicativo. Contudo, eles devem descrever com detalhes as situações de uso que refletem um modelo de tarefas humanas, representando assim, os requisitos do usuário. Normalmente são implementados entre as fases iniciais de análise e de especificação da interface. Também podem auxiliar os projetistas a obter uma visão global, de todos os processos comunicativos, existentes entre o usuário e o aplicativo.

Ao construir um cenário, o projetista deve saber qual é o objetivo a ser atingido entre esses dois agentes, neste processo de comunicação. Alguns fatores de domínio e sócio-culturais, também devem ser levados em consideração, no momento da criação desses cenários. Bem como, propor a estruturação das tarefas ou uma solução de interação coerente.

As dúvidas quanto ao desenho do projeto e os questionamentos identificados, durante a construção dos cenários, deverão ser esclarecidos através de um conjunto de perguntas definidas pelo projetista. Os cenários, também permitem ao projetista identificar os signos que farão parte da aplicação, desde os signos que fazem parte do domínio, até os que surgiram pelo processo de informatização das tarefas do usuário.

### 3.3.1 Signos

Os signos citados por Silva e Barbosa (2004) são identificados na interface como objetos de interação e podem ser obtidos durante o processo de criação dos cenários. As tarefas que foram identificadas nos cenários, alteram ou apresentam algum tipo de informação e é justamente neste processo comunicativo, executado pelo usuário, que surge o signo. Os signos podem ser classificados como sendo de domínio, transformados e de aplicação.

- **Signos de Domínio:** São facilmente identificados, dada a sua familiaridade com os

objetos da vida real. Normalmente, não necessitam de maiores explicações, ao menos, que tenham alguma característica muito particular. Como por exemplo: O endereço do usuário é algo muito familiar para ele. Porém, se houver alguma particularidade, quanto ao preenchimento do seu conteúdo, será necessário uma melhor elucidação.

- **Signos Transformados:** São signos que originalmente tem características de um signo de domínio, porém passam por algum tipo de alteração de escopo dentro da aplicação. Podem apresentar características para criação de analogias e metáforas. Como por exemplo: O nome do usuário é um signo de domínio. Porém, pode ser utilizado para autenticação do usuário, através de um processo biométrico de reconhecimento da fala.
- **Signos de Aplicação:** São identificados pela característica particular, de apenas fazerem sentido dentro da aplicação, não tem nenhum significado para o usuário. Portanto devem ser esclarecidos. Como por exemplo: A funcionalidade do conjunto de opções, denominado atalho, deve ser totalmente esclarecido ao usuário, para que possa ser utilizado adequadamente.

A classificação coerente dos signos, permite que o projetista possa proceder corretamente no momento da construção dos objetos de interação. Os signos podem ser agrupados respeitando algum critério que seja coerente para as regras de negócio do aplicativo. Como por exemplo: Os signos simples RENAVAL, placa, cor, chassi, marca e modelo, podem ser agrupados em um único signo composto chamado veículo.

Os signos podem ser representados em uma tabela conforme exemplo abaixo. Ela também poderá ser alterada, de acordo com a evolução do processo de modelagem.

Signo	Definição	Tipo
RENAVAM	código RENAVAL	Aplicação
Placa	Número da placa do veículo	Domínio
ano de fabricação	data de fabricação do veículo	Domínio
cor	cor do veículo	Domínio
Tipo	tipo do veículo	Aplicação
Chassis	número do chassis do veículo	Domínio
marca	marca do veículo	Domínio
Modelo	modelo do veículo	domínio
combustível	tipo de combustível do veículo	Domínio

## Tabela 1 – Signo composto - veículo

### 3.3.2 Perguntas

As perguntas na MoLIC, conforme citado por Silva e Barbosa (2004) tem o objetivo de complementar algumas falhas de interpretação do projetista, que podem ter sido geradas ao longo da criação dos cenários. Situações de ambigüidade ou dúvidas poderam ser sanadas através do seu uso.

Os projetistas podem criar um conjunto de perguntas que deverão ser associadas aos cenários. Essas perguntas, deverão ser identificadas através de um número entre colchetes, dentro do texto do cenário que corresponde a sua abordagem. O exemplo a seguir ilustra os procedimentos acima relatados.

#### **Perguntas exploradas nos cenários**

[1] Quem irá utilizar o serviço do Portal de Voz Governamental?

[2] Quais serão os serviços fornecidos pelo Portal de Voz Governamental?

O termo PVG utilizado nos cenários, significa Portal de Voz Governamental.

#### Cenário 1- Informações sobre o IPVA

João ficou sabendo, através do seu amigo Arnaldo, da existência de um tal de portal de voz que fornece serviços governamentais para qualquer cidadão [1]. Ele gostaria de utilizá-lo, porém somente foi lhe dado um número de um telefone e ele não tem a mínima idéia de como proceder para obter determinado serviço.

“Bem, já que estamos no começo do ano, gostaria de sanar algumas dúvidas que tenho a respeito do IPVA.”

Depois de alguns dias João se encontrou com o Arnaldo e contou a sua experiência da seguinte forma:

João: Telefonei para aquele número que você me passou e os nossos processos interativos foram os seguintes:

PVG: *“mensagem de boas vindas e optar pelo serviço desejado”* [2]

João: “ ... “

Cabe observar que nos cenários, as perguntas que são representadas por números entre colchetes, respondem as dúvidas ou os questionamentos que o projetista deseja esclarecer ou ressaltar durante o processo interativo.

### 3.4 Modelo de Tarefa

O modelo de tarefas, conforme citado por Silva e Barbosa (2004) é a próxima etapa a ser executada dentro de um projeto de IHC, baseado nos conceitos da MoLIC. Inicialmente, o

objetivo é o de obter, através dos cenários, subsídios para identificar as principais metas do usuário ao utilizar o aplicativo. Cabe ressaltar, que uma meta pode estar contida em mais que um cenário e um cenário pode descrever mais que uma meta.

### 3.4.1 Diagrama Hierárquico de Metas

O diagrama hierárquico de metas deve conter o conjunto de metas identificadas nos cenários. Esse diagrama fornece uma visão global das metas do usuário ao utilizar o aplicativo e deverá ser organizado segundo um critério pré-definido pelo projetista.

A representação da meta nesse diagrama, contém o nome da meta sob a perspectiva do usuário. Esse nome pode estar baseado em uma fala do usuário como: "Eu quero utilizar o aplicativo para <meta>". Essa representação, também é composta pelos papéis que o usuário executa para atingir essa meta e deve conter uma letra para a sua identificação dentro do diagrama.

Os critérios para classificação das metas devem ser definidos pelo projetista. Questionamentos como que metas devem ser priorizadas, que metas demandam mais tempo para serem modeladas, que metas se relacionam entre si e outras. São perguntas que o projetista deve fazer para tentar definir o critério de classificação das metas.

Com o objetivo de exemplificar essas afirmações e facilitar o entendimento, a seguir serão apresentados os critérios de classificação das metas escolhidos para o portal de voz governamental:

1. Utilização do portal
2. Suporte ao portal

O diagrama hierárquico de metas no estudo em questão, apresentará a estrutura da árvore utilização do portal, que é um dos dois critérios de classificação de metas escolhido para representar o portal de voz governamental. A sua demonstração é descrita em forma textual. Um exemplo do diagrama hierárquico de metas para o uso do portal está a seguir.

1. Utilizar o portal
  - 1.1. consultar IPVA
    - 1.1.1 Valor do IPVA - B

- 1.1.2 Saldo Devedor - C
- 1.2. pagar IPVA
  - 1.2.1 Ano Corrente - D
  - 1.2.2 Anos Anteriores - E

A cada meta identificada no diagrama deve ser associado um modelo de tarefas que é composto de uma decomposição hierárquica dos passos necessários para atingir essa meta, sob a perspectiva do usuário. Sendo assim, cada tarefa pode ser decomposta em subtarefas, e cada subtarefa pode ser novamente decomposta em novas subtarefas.

### 3.4.2 Estruturas de Tarefas

Os tipos de estruturas existentes para organizar as tarefas são: seqüenciais, independentes de ordem, alternativas e iterativas.

A estrutura seqüencial prioriza a ordem que as tarefas necessariamente devem ser executadas pelo usuário. Já na estrutura de tarefas independente de ordem as tarefas podem ser executadas em qualquer ordem, pois representam um conjunto de tarefas não seqüenciais a serem executadas pelo usuário. Neste caso o projetista pode sugerir uma seqüência, porém será o usuário que realmente irá definir essa ordem de execução.

A estrutura alternativa permite que o usuário defina qual tarefa será executada para se atingir determinada meta. Neste caso, o usuário tem total liberdade para definir a ordem de execução das tarefas. Já a estrutura iterativa é caracterizada pela repetição de uma tarefa para se atingir uma determinada meta. As expressões abaixo indicam a cardinalidade da execução de uma tarefa:

[n+] indica que a tarefa deve ser realizada pelo menos n vezes;

[m..n] indica que a tarefa deve ser realizada no mínimo m e no máximo n vezes.

### 3.4.3 Tipos de Tarefas

Na MoLIC, conforme citado por Silva e Barbosa (2004), existe um grupo de tarefas especiais, que foram divididas em tarefas opcionais e ubíquas. Elas serão descritas com maiores detalhes a seguir:



As tarefas opcionais ocorrem quando o usuário pode optar por realizar ou não uma tarefa. As tarefas ubíquas podem ser feitas em qualquer ponto da realização da meta. Por exemplo: no portal de voz governamental o processo de abandonar opção, sair do portal ou falar com o atendente foi definida como tarefa ubíqua. Dessa maneira, a interface deve prover ao usuário a execução dessa tarefa, independente da solução tecnológica escolhida.

#### 3.4.4 Pré-condições

Na execução de algumas tarefas ou metas é comum existir pré-condições, que são estados que se devem garantir para a execução da tarefa. Como por exemplo: No portal de voz governamental a tarefa efetuar login é pré-condição para a tarefa pagar IPVA.

#### 3.4.5 Re-uso de Tarefas e Metas

As tarefas que fazem parte da estrutura de várias metas são definidas como estereótipos. Um estereótipo de tarefa pode ou não receber parâmetros e podem ser representados da seguinte forma:

<<nome\_do\_estereótipo(parâmetro1, parâmetro2, ...)>> ou <<nome\_do\_estereótipo>> (no caso de não haver parâmetros).

No portal de voz governamental a tarefa efetuar login, foi também definida como um estereótipo.

Outra maneira de re-uso é referenciar diretamente as metas ou tarefas já especificadas dentro desta mesma modelagem, porém não suficientemente genéricas para se tornarem um estereótipo. Para as metas basta incluir a expressão Nome da Tarefa = META X, já nas tarefas, é utilizada a expressão Nome da Tarefa = X.Y e nas subtarefas, a expressão é Nome da Tarefa = X.Y.Z. Onde:

X representa a meta ;

Y a tarefa;

Z a subtarefa.

#### 3.4.6 Representação dos Signos no Modelo de Tarefas

A MoLIC, por implementar os conceitos da Engenharia Semiótica, reforça a necessidade de

representar os signos e as falhas de comunicação ao longo de um processo interativo. Sendo assim, a representação diagramática - que é composta por uma especificação textual, que visa fornecer uma visão geral das metas e da hierarquia das tarefas - é complementada pela identificação dos signos existentes em cada tarefa e pelo tratamento de erros previstos pelo projetista nesta fase da modelagem.

Os cenários exercem um papel fundamental no processo de definição dos signos, pois auxiliam na construção da tabela de signos e na identificação de como, quando e onde os signos são utilizados no processo interativo. A etapa posterior a modelagem diagramática das tarefas, consiste em associar as tarefas aos signos correspondentes, e na especificação textual, informar a palavra-chave SIGNO e incluir os signos identificados.

Para referenciar um signo que está contido em um signo composto, deve se proceder da seguinte forma nome\_signo\_composto.nome\_signo. Quando for necessário fazer referência a todos os signos de um determinado signo composto, deve se proceder da seguinte forma nome\_signo\_composto.\*. Por exemplo, o signo correspondente a placa do veículo pode ser expresso como veículo.placa, enquanto em uma situação em que são apresentados ou manipulados todos os dados de um veículo, o conjunto destes signos pode ser representado por veículo.\*.

Quando o signo for utilizado para apresentar o seu conteúdo (saída) é representado por seu nome seguido de ponto de exclamação (signo1!). Já um signo que corresponde a uma informação fornecida pelo usuário (entrada) é representado por seu nome seguido de ponto de interrogação (signo2?). Se houver várias instâncias do mesmo tipo a representação é (nome\_do\_signo,cardinalidade).

#### 3.4.7 Prevenção e Tratamento de Erro

Os problemas na comunicação ao longo da interação, que podem ser identificados previamente pelo projetista, no momento da criação do modelo de tarefas, devem ser tratados em cada tarefa. Cada um deles tem a sua característica particular e estão classificados segundo as seguintes categorias:

a) prevenção passiva: Prevenção de erros através de instruções durante o uso do aplicativo ou

por documentação prévia. Por exemplo: "O usuário não possui privilégio para acessar o sistema".

b) prevenção ativa: Prevenção de erros através da identificação dinâmica, ao longo da utilização do aplicativo. Por exemplo: tarefas que devem ser habilitadas ou desabilitadas em determinadas situações.

c) prevenção apoiada: Prevenção de erros que o aplicativo identifica, mas solicita uma confirmação por parte do usuário. São situações que o sistema consegue detectar como sendo erros em potencial, mas cuja decisão recai sobre o usuário. Normalmente são caracterizados por utilizar mensagens de confirmação. Por exemplo: "Arquivo existente, deseja substituir?"

d) tratamento apoiado: Prevenção de erro que deve ser tratada pelo usuário, com suporte do preposto do projetista. Por exemplo: Informar ao usuário da existência de um erro e fornecer a opção de corrigi-lo.

e) captura de erro: Erros identificados pelo aplicativo, que são informados ao usuário, mas não existe ação possível dentro do aplicativo. O seu objetivo é o de apenas informar o fato. Por exemplo: "Existe um arquivo danificado".

### 3.4.8 Estrutura Textual Completa do Modelo de Tarefas

Na MoLIC conforme citado SILVA e BARBOSA (2004) o modelo de tarefas completo é composto pelos cenários associados a meta a ser atingida, os papéis dos usuários e a seqüência de tarefas.

Com o objetivo de exemplificar esse modelo, abaixo está descrito o pagamento do ano corrente do IPVA, utilizado no portal de voz governamental.

Meta D: Pagar ano corrente do IPVA

Cenários Associados: pagamento do IPVA e IPVA para veículo novo

Papéis: Usuário Cadastrado e Usuário Proprietário

D.1. Fornecer Identificador Único do Veículo

SIGNOS: veículo.RENAVAM?

PREVENÇÃO ATIVA: somente usuários cadastrados ou proprietários de veículo podem executar esse tipo de consulta

TRATAMENTO APOIADO: falta de informações obrigatórias

D.1.a. Tipo de Pagamento

D.1.a.1. Único com desconto

SIGNOS: ipva.valor, 0..1!

D.1.a.2. Único sem desconto

SIGNOS: ipva.valor, 0..1!

D.1.a.3. Parcelado

SIGNOS: conjunto(ipva.data\_1, ipva.valor\_1, ipva.data\_2, ipva.valor\_2, ipva.data\_3, ipva.valor\_3, 0..1)!

D.2. Forma de pagamento

D.2.a. Imediato

SIGNOS: banco.banco?, banco.agência?, banco.conta?, banco.senha?

TRATAMENTO APOIADO: falta de informações obrigatórias

D.2.b. Posterior

D.3. Emitir Comprovante

D.3.a.1. Por e-mail

SIGNOS: usuario.e-mail?

D.3.a.2. Por fax

SIGNOS: usuario.fax?

D.3.a.3. Por correio

SIGNOS: usuario.endereço?, usuario.cep?

### 3.5 Modelo de Interação

Conforme cita Paula(2003), a MoLIC recomenda que o projetista, crie o modelo de interação de uma forma bastante minuciosa, baseado no modelo de tarefas, antes de definir a especificação da interface. A Engenharia Semiótica trata a interação, como sendo um processo de comunicação realizado entre o usuário e o preposto, que é o representante do projetista dentro da interface. Esta conversa entre esses dois elementos é fundamental no processo interativo e deve ser encarada com muita seriedade. As ferramentas de desenvolvimento de IHC podem ajudar, e muito, o projetista a identificar todas as conversas existentes entre o usuário e o aplicativo.

A MoLIC conforme cita Paula(2003) propõe um modelo diagramático de interação, em conjunto com as especificações textuais, para modelar essas conversas. O modelo de interação fornece uma visão geral de todos os processos interativos ocorridos.

As cenas, os processos do sistema e as transições, são os elementos que compõem o diagrama de interação. A cena aborda uma conversa sobre determinado assunto entre o usuário e o preposto, que geram pares conversacionais. Uma cena pode ser associada, a uma tela ou a uma página, no modelo de interface.

Uma cena, conforme Silva e Barbosa (2004) é composta por um identificador do tópico, ou assunto, descrito por um verbo no modo infinitivo, sob a perspectiva do usuário, e por um conjunto de diálogos, realizados entre o usuário e o preposto, que devem estar entre colchetes.

Um processo, conforme Silva e Barbosa (2004) é caracterizado por ser uma forma de comunicação do preposto com o usuário final. Ele é a única forma que o projetista tem para informar o usuário do resultado de um determinado processamento do sistema. Os processos somente necessitam ser criados quando o seu resultado precisa ser comunicado ao usuário, isto é, dependem da intervenção do usuário para continuar o processamento.

Uma transição, conforme Silva e Barbosa (2004) representa uma alteração no fluxo da conversa, normalmente provocado por uma fala do usuário (cena) ou por uma fala do preposto (processo). Uma transição é composta por um rótulo que contém três elementos, a saber:

a) pré-condições:

condições que devem ser previamente atendidas para que o usuário ou o preposto possam iniciar o processo comunicativo. São representados pela palavra-chave PRÉ:. Podem ser informados através do uso de uma linguagem natural ou por uma expressão lógica. Por exemplo, pré: O usuário necessita estar cadastrado.

b) Falas do usuário ou do preposto do projetista.

As transições provocadas pelas falas do usuário são expressas pela palavra-chave *u:[fala do usuário]*. Por exemplo, *u:[confirmar exclusão]*. Já as transições provocadas pelas falas do preposto são expressas pela palavra-chave *p:fala do preposto*. Por exemplo, *p: erro - usuário*

*não cadastrado.*

c) pós-condições:

condições que passam a ser verdadeiras durante a transição. São expressas pela palavra-chave *pós:*. Por exemplo, *pós: exclusão efetuada.*

### 3.5.1 Modelo Diagramático de Interações

O modelo diagramático de interação tem como principal objetivo, fornecer uma visão global aos projetistas do discurso interativo. . Para facilitar a compreensão e esclarecer as etapas desse modelo, será descrita a interação efetuar login, executada no portal de voz governamental.

Cena: Efetuar login

pré: login = Falso

1. u: [fornecer dados de login]  
 p: erro - login inexistente  
 p: repetir ou adicionar login?

1.1. u: [repetir]  
 pós: login = Falso (retornar à cena efetuar login)

1.2 u: [adicionar login]  
 pós: login = Falso (cena adicionar usuário)

2. u: [fornecer dados de login]  
 p: sucesso  
 pós: login = Verdadeiro (retornar à cena principal)

3. u: [abandonar login]  
 pós: login = Falso (cena qualquer)

### 3.5.2 Especificação Textual

A especificação textual fornece os detalhes de cada conversa existente no modelo de interação. O mesmo exemplo será utilizado neste caso. A saber:

Cena: Efetuar login

1. Fornecer dados do login  
nome? <nome completo do usuário: obrigatório>

### 3.6 Considerações finais

Esta seção apresentou os principais conceitos de uma linguagem de modelagem para sistemas interativos, em especial, a MoLIC que através dos seus artefatos, fornece subsídios para que o projetista possa criar aplicativos que favoreçam a interação com o usuário final. A próxima seção fará um levantamento das principais funcionalidades do portal atual do governo do Estado de São Paulo e fornecerá propostas para a criação de um portal de voz. Também, fará uma análise do conjunto de soluções existentes no mercado, nacional para os portais de voz, destacando os processos, linguagens e plataformas de desenvolvimento dessa tecnologia.

#### 4. PORTAL DE VOZ COM FOCO EM GOVERNO ELETRÔNICO

O fato de já existirem vários serviços disponíveis aos cidadãos do Estado de São Paulo, em um único sítio de governo eletrônico <http://www.cidadao.sp.gov.br>, citado em São Paulo (2006), é um fator importante a ser considerado na construção de um portal de voz, com esses mesmos fins. Isso significa que, a princípio, nada será criado, apenas haverá uma mudança na forma de interação do usuário com o aplicativo, para se obter o serviço desejado. Inicialmente, será apresentado o modelo arquitetônico de um portal de voz, conforme a Figura 1, onde são demonstrados os principais componentes que fazem parte da solução.

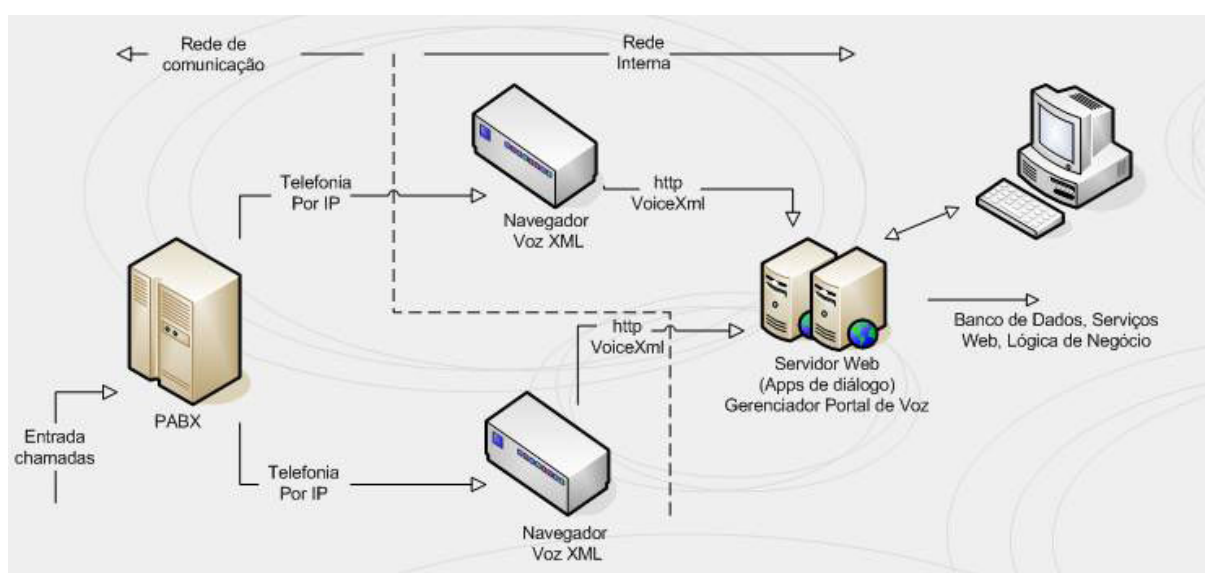


Figura 1 – Arquitetura de rede de um portal de voz e componentes de software

É importante ressaltar, que a proposta desse estudo não é a de comparar a contribuição que ambos os portais trazem para a população em geral, mas sim, propor uma outra alternativa de acesso aos serviços que o poder público deve fornecer ao cidadão do estado de São Paulo. Sendo assim, o portal de voz surge como uma nova alternativa, mais democrática e acessível, para obter essas informações.

Esta seção visa discutir o processo de construção do portal de voz: os problemas encontrados durante o processo de migração e as suas possíveis soluções, inclusive do ponto de vista da tecnologia para implementá-lo.



#### 4.1 Migração do Sítio de Governo Eletrônico para um Portal de Voz

O processo de adequação do sítio atual para a tecnologia de um portal de voz não é uma tarefa trivial. A idéia inicial seria a de se aproveitar a estrutura já existente e migrá-la para essa nova tecnologia. A principal intenção desse procedimento será o de se evitar o retrabalho nos itens que atendem plenamente os conceitos de acessibilidade e de usabilidade.

É óbvio que haverá alteração nos processos de interação, porém a infra-estrutura atualmente existente para o armazenamento dos dados será mantida.

Portanto, o objetivo será o de adaptar os serviços que hoje estão nos sítios de governo eletrônico para um portal de voz. Para tal, torna-se necessário analisar alguns aspectos das aplicações visuais e adequá-las para as características das aplicações audíveis. Alguns dos itens a serem observados é o padrão, a funcionalidade, a arquitetura e o uso.

##### 4.1.1 Quanto ao Padrão

O portal do cidadão é o sítio oficial do Governo do Estado de São Paulo para prestar serviços à população. Nele são encontrados alguns aspectos interessantes, como processos interativos com foco nas atividades da vida diária que facilita, e muito, a comunicação do aplicativo com o usuário. Outro ponto favorável é que ele contém vários serviços e informações aos cidadãos; bem como ligações com os diversos sítios das Secretarias de Estado, o que o torna bastante abrangente. É notório que é uma construção que privilegia os aspectos visuais, não sendo uma exceção aos sítios existentes na internet, o que é óbvio, pois os tornam altamente intuitivos.

Neste sítio existe uma grande quantidade de serviços associados às várias Secretarias de Estado e outros sítios com o mesmo tipo de proposta, o que o torna muito abrangente, no que se refere à variedade de serviços prestados aos seus visitantes. Entretanto, cada um desses sítios possui o seu padrão próprio de desenvolvimento, no que se refere ao desenho, nomenclatura, formato, entre outras características, fazendo com que não haja uma consistência entre as páginas acessadas. Esta inconsistência é derivada do processo de desenvolvimento dos sítios. A responsabilidade pelo desenvolvimento é dos vários órgãos do

Governo do Estado, e não há requisitos que estes sítios tenham que apresentar consistência em seus processos de navegação e organização de conteúdo.

Já no portal de voz proposto, pretende-se que haja uma maior padronização, pois o sistema será totalmente construído baseado em uma padronização bem definida para sua organização (árvore de opções), textos das opções (fraseologia), rapidez no acesso (atalhos) e formatação (tipo de voz).

#### 4.1.2 Quanto à Funcionalidade

Conforme dito anteriormente, a funcionalidade do novo portal deverá ser a mesma do portal visual. A diferença concentra-se na forma de interação, como descrito a seguir.

O portal de voz apresentará o retorno das informações desejadas, por meio da comunicação verbal, ou em alguns casos, pela emissão de um documento da operação realizada, através do serviço de correio convencional, por fax, ou mesmo, por correio eletrônico. No portal visual, o usuário pode ler na tela as informações desejadas e imprimi-las em uma impressora local, obtendo de forma imediata esse documento.

Essas diferenças, na sua essência, não se caracterizam como algo relevante, no que se refere a funcionalidade do aplicativo em uma análise comparativa entre as duas soluções. Contudo, cabe observar que a forma de interação para acessar essas informações é completamente diferente.

#### 4.1.3 Quanto ao Desenho

O princípio fundamental de ambos os portais é o mesmo, isto é, o atendimento as necessidades do cidadão, no que se refere as informações e serviços do Estado, pertinentes a todos os contribuintes, proporcionando a eles um acesso rápido e seguro de forma eletrônica. Para que essa análise não incida em erros graves de interpretação, torna-se necessário que haja uma divisão quanto a alguns aspectos que individualizam e particularizam cada um desses sítios.

Inicialmente, eles podem ser analisados com base nos fatores tecnológicos de cada plataforma de desenvolvimento. No portal visual, o foco é nos aspectos visuais como a principal forma de interação entre o usuário e o aplicativo. As imagens e os ícones possibilitam uma identificação relativamente fácil dos serviços, através do uso da intuição e da associação dos usuários com a imagem em questão. Já o portal proposto deve privilegiar a objetividade e a clareza na interação, pois a única forma de comunicação entre o usuário e o aplicativo é a fala humana. O desenho dessa interface tem o objetivo principal de permitir o fácil acesso aos serviços desejados, sem provocar o cansaço e a dispersão do usuário.

A observação das heurísticas para os portais de voz, nessa etapa do desenvolvimento é de fundamental importância. Itens como o balanceamento da árvore de opções, fraseologia, suporte ao usuário e flexibilidade, devem ser implementados e respeitados na sua íntegra.

#### 4.1.4 Quanto ao uso

Existem várias diferenças na utilização desses portais, mas a principal, e a que cabe ser ressaltada é o aspecto de inclusão digital que no portal de voz é observado. Ele permite um fácil acesso a todos os seus serviços, notoriamente as pessoas analfabetas, as idosas e as portadoras de necessidades especiais ou não. Essa sua flexibilidade e abrangência tornam esse portal, um enorme agente facilitador para os cidadãos utilizarem esse tipo de serviço.

Outro aspecto relevante é o processo de interação e navegação entre os serviços. O portal visual permite métodos empíricos de acesso, de tentativa e erro, possibilitando idas e vindas nas diversas opções. Já no portal de voz, este tipo de interação é péssima, pois provoca ao usuário cansaço e frustração; bem como indisponibilidade no serviço devido à longa duração da ligação telefônica.

## 4.2 Problemas e Soluções Durante o Processo de Migração

A conversão do portal visual para um portal de voz representa além de uma mudança tecnológica, uma mudança na filosofia de desenvolvimento de um novo sistema.

Preocupações como o formato dos diálogos interativos, objetividade nas opções, clareza nas mensagens trocadas entre o aplicativo e o usuário e quantidade de opções em uma única interação, passam a ser extremamente importantes para as interfaces audíveis.

Situações de sobrecarga de informações e falta de clareza nos diálogos, que normalmente passam despercebidas em um aplicativo visual, devem ser tratadas com o maior cuidado nas aplicações audíveis, pois o item imagem não faz parte nesse processo de comunicação. Portanto, esse tipo de interface deve conter diálogos muito claros e objetivos entre essas duas entidades.

O Governo do Estado de São Paulo tem uma enorme quantidade de serviços a serem oferecidos aos cidadãos do Estado. No portal visual, essa característica não é muito crítica, pois em uma única interação, dentro de uma interface visual, pode-se transmitir uma grande quantidade de informações, através de imagens, sem prejudicar o processo de comunicação. Porém, no portal de voz, essa particularidade deve ser tratada com muita seriedade. Neste tipo de aplicativo, por ser audível, não existe a facilidade de transmitir essa mesma quantidade de informações em um único diálogo, sem sobrecarregar o usuário.

Como alternativa para este problema, a proposta do estudo é a de utilizar no primeiro nível de interação, uma nomenclatura mais familiar ao usuário como o nome do serviço desejado ou algo similar ao portal do cidadão, como eventos da vida diária, ao invés de uma nomenclatura mais formal, como o nome da Secretaria de Estado que fornece este tipo de serviço.

A criação da árvore de opções para o portal de voz, serve como um excelente exemplo para reforçar essa afirmação. Se durante o processo de construção dos diálogos, a opção for a de utilizar os nomes das Secretarias de Estado, como Fazenda, Planejamento, Saúde, Educação, Transporte, entre outras, é muito provável que o usuário tenha dificuldade em encontrar o serviço desejado. Contudo, se a opção for para algo mais familiar para o usuário como impostos, hospitais, escolas, polícia, luz, água e esgoto, esse acesso as informações será mais simples e direto. Apesar desse primeiro nível de interação diminuir a abrangência de cada uma das opções, este aspecto negativo será atenuado, com a inclusão de um item no menu, denominado secretarias, onde em um segundo nível de interação, para usuários mais experientes e com uma maior abrangência nos serviços, o usuário acessará a secretaria de estado desejada.

Baseado nessas afirmações e nas informações já fornecidas ao longo dessa seção, pode-se verificar que a adaptação do portal visual para o portal de voz, é uma tarefa muito complexa. Para levá-la a bom termo, somente a implementação dos conceitos de IHC, principalmente a execução de testes de usabilidade na sua concepção, poderão minimizar os seus impactos. Cabe lembrar, que a fraseologia em um portal de voz exerce um papel fundamental na comunicação entre esses dois elementos.

A utilização de uma modelagem da interação, no caso usando a MoLIC provou ser um grande aliado para descobrir os objetos de diálogo e definir os principais requisitos de interação, necessários para a construção de processos comunicativos adequados.

Outro elemento de grande utilidade neste processo são as heurísticas para os portais de voz que, com o seu conjunto de sugestões de boas práticas, forneceram subsídios para solução de alguns dos problemas enfrentados.

O processo de uso da MoLIC para o projeto da interação seguiu os passos:

1. Foram criados cenários e seus elementos derivados, como as perguntas e os signos. Os cenários auxiliaram a descobrir os principais diálogos interativos entre o aplicativo e o usuário.
2. Subsidiado pelo resultado do passo anterior, definiu-se o próximo elemento da modelagem, que é o modelo de tarefas do usuário. Nessa etapa, após analisar detalhadamente qual era o objetivo do usuário ao executar determinadas tarefas, foi possível identificar as metas do usuário.
3. Por fim, o modelo de interação forneceu uma visão geral de todos os processos interativos existentes no aplicativo, possibilitando obter uma noção global das inter-relações entre as tarefas identificadas na etapa anterior da modelagem.

Durante todo o processo, as heurísticas para os portais de voz auxiliaram no sentido de orientar quais seriam as melhores práticas a serem aplicadas no desenvolvimento do modelo do aplicativo:

- a heurística de balanceamento da árvore de opções foi um instrumento útil para determinar um equilíbrio nos diálogos interativos.

- A heurística de fraseologia contribuiu para tornar a comunicação mais clara e objetiva. Privilegiou-se o uso da linguagem natural durante o processo comunicativo.
- A heurística de suporte ao usuário contribuiu para se identificar a necessidade da criação de processos de ajuda e saída do sistema.
- A de flexibilidade possibilitou formas de acesso mais rápidas às opções desejadas.
- Por fim, a heurística de confiabilidade ajudou no sentido de tornar a interação mais estável e robusta.

Compreende-se que existe a necessidade de melhorar esse modelo, mesmo porque ele é uma primeira versão do processo de modelagem do portal de voz com foco em governo eletrônico. Como todo processo visando qualidade e mais especificamente, como processo de desenvolvimento centrado no usuário, a modelagem é sujeita à melhoria continuada. Sugere-se que técnicas de avaliação de usabilidade sejam usadas para aprimorar esses processos interativos.

#### 4.3 Linguagens e Padrões da Tecnologia de Portais de Voz

Após a conclusão das etapas conceituais do projeto de construção de um portal de voz, virá a fase de implantação. Nesse momento, as informações fornecidas ao longo dessa seção, serão de grande utilidade para auxiliar a tomada de decisão, sobre qual é a plataforma de desenvolvimento mais adequada. Esse estudo não pretende implementar a solução de um portal de voz governamental, mas se propõe a fornecer subsídios para facilitar esse processo.

Atualmente, o mercado de desenvolvimento de sistemas para reconhecimento de voz, conta com duas linguagens principais para construir suas aplicações. O VoiceXML ou o VXML (*Voice Extensible Markup Language*) e o SALT (*Speech Language Application Tags*), cada uma delas com as suas particularidades. Ambas linguagens podem oferecer vários recursos para o desenvolvimento de aplicações baseadas em comandos de voz.

Nesta seção serão apresentadas as características mais importantes que as individualizam,

bem como pretende fornecer uma visão geral de cada uma dessas linguagens, e os padrões que as norteiam, durante o processo de desenvolvimento de sistemas que utilizam esse tipo de tecnologia.

Inicialmente serão apresentados os aspectos gerais de ambas as linguagens, VoiceXML e SALT, para que possa ser fornecida uma visão geral do funcionamento de cada uma delas. Posteriormente, o foco será o de fazer uma análise comparativa entre as duas linguagens com base nessas características. Cabe ressaltar que todas as informações foram obtidas, nos sítios oficiais do Forum de cada um dos padrões de linguagem, através dos documentos do *Voice extensible Markup Language (2.0 Specification)*[39] citado em WORLD WIDE Web Consortium (2004) e do *Speech Application Language Tags (1.0 Specification)* citado por CISCO Systems Inc. et al.(2002). E por fim, fornecer algumas vantagens e desvantagens, recomendações de uso e quais os fatores que o projetista deve levar em consideração no momento de definir qual a linguagem deverá ser utilizada durante o processo de desenvolvimento de uma aplicação de voz.

#### 4.3.1 VoiceXML ou VXML

Segundo informações obtidas pelo Forum VXML, o VoiceXML teve como origem de seu desenvolvimento, os laboratórios da AT&T Bell, em um projeto denominado *PhoneWeb*. Após a separação da AT&T e Lucent, ambas procuraram desenvolver de forma independente, suas próprias versões, de uma linguagem com esses mesmos fins. Os Laboratórios da Lucent Bell continuaram trabalhando em um projeto, conhecido como TelePortal. Onde o seu principal foco foi o de criar serviços e aplicações com uso de linguagem natural. Já os laboratórios da AT&T desenvolveram uma linguagem e uma plataforma que tem sido utilizada para construir vários tipos de aplicações, que compreendem centros de atendimento, serviços telefônicos ao consumidor e serviços, através da internet, que o próprio usuário configura e administra as características do seu telefone.

O amadurecimento dessas várias linhas de pesquisa, resultou no desenvolvimento do padrão VoiceXML, e em seguida, a Motorola aderiu à linguagem para prover informações de Web aos usuários proprietários de telefones móveis. Em outubro de 1998 foi feito o primeiro anúncio da tecnologia. Outras empresas se interessaram pelo conceito e foi fundado o

VoiceXML Forum, que tem como patrocinadoras a AT&T, IBM, Lucent e Motorola. Inicialmente, os principais objetivos eram:

- a) Desenvolver uma especificação VXML aberta e posteriormente torná-la um padrão de linguagem;
- b) Conscientizar as empresas da necessidade da existência de uma linguagem padrão para o desenvolvimento de aplicações de voz;
- c) Convencer as empresas a apoiarem e participarem do VoiceXML Forum;
- d) Incentivar as empresas a utilizar esse padrão no desenvolvimento de aplicações e serviços de voz.

Atualmente, cerca de 600 empresas participam do grupo, que tem como missão estabelecer novas especificações e promover no mundo o padrão dessa linguagem. A especificação 0.9 foi liberada em Agosto de 1999, e algumas empresas começaram a utilizá-la em seus produtos e serviços. Já em Março de 2000 a especificação 1.0 foi lançada, baseada nos últimos anos de pesquisa e desenvolvimento de aplicativos pelas empresas que criaram o padrão, e também, com os comentários enviados por mais de 150 empresas que participam do fórum. Em Outubro de 2001 o padrão foi adotado pelo W3C com a versão 2.0 e, desde então, sua utilização pelas empresas que desenvolvem este tipo de aplicação vem crescendo de forma significativa.

O VoiceXML foi projetado para criar diálogos auditivos com características de sintetização da fala humana, áudios digitalizados, reconhecimento de voz e entrada através de dispositivos como teclas DTMF, gravação de mensagens faladas, telefonia, e conversações heterogêneas. Seu principal objetivo é o de trazer as vantagens de um desenvolvimento baseado nos padrões da internet e permitir um diálogo de forma interativa em aplicações com reconhecimento de voz. O VoiceXML foi baseado no XML, que garante uma boa estruturação dos códigos e proporciona uma documentação adequada. Bem como, fornece um conjunto de elementos declarativos de um diálogo, como por exemplo: "*prompts*", eventos, gramáticas, entre outros. Essas características tornam o padrão de linguagem VoiceXML uma ferramenta bastante completa.

Segundo o documento de especificação VoiceXML citado em WORLD Wide Web Consortium (2004) o seu principal objetivo é o de oferecer aos projetistas uma plataforma



completa de desenvolvimento e entrega de conteúdo para aplicações que utilizam o reconhecimento de voz como forma de interação, e proporcionar aos desenvolvedores desse tipo de aplicação uma interface de alto nível, o que facilita consideravelmente o seu desenvolvimento. Também permite integração de serviços de voz com serviços de dados de forma amigável.

O VoiceXML está baseado nos mesmos princípios de desenvolvimento da linguagem XML, a saber:

- a) A linguagem permite portabilidade de serviços através da abstração de recursos entre as plataformas;
- b) A linguagem aceita uma grande diversidade de plataformas, suportando vários formatos, como arquivos de áudio, gramáticas de fala e URIs;
- c) A linguagem permite ao desenvolvedor facilmente implementar em seu projeto os tipos de interação mais comuns;
- d) A linguagem tem uma boa definição semântica, que preserva a intenção do desenvolvedor, considerando as condições de interação com o usuário. Heurísticas do cliente não são necessárias para a interpretação dos elementos do documento;
- e) A linguagem tem um mecanismo de controle de fluxo;
- f) A linguagem permite uma separação dos serviços lógicos das condições de interação;

#### 4.3.2 SALT

O padrão dessa linguagem foi elaborado e concebido, após um grupo de empresas se reunirem e fundarem o SALT Forum. Entre os principais participantes dessa organização estão a Cisco, Converse, Intel, Microsoft, Phillips e Speech Works. O SALT atualmente se encontra na versão 1.0, baseada na especificação de Julho de 2002.

O SALT é uma linguagem de marcação para o desenvolvimento de interfaces de voz e consiste principalmente em um pequeno conjunto de elementos XML, que associam atributos e propriedades de objetos DOM (*Document Object Model*), eventos e métodos, que implementam uma interface de fala para páginas da internet. O SALT pode ser usado em conjunto com HTML, XHTML e outros padrões para construir interfaces de fala, como sistemas que utilizam apenas a voz como forma de interação e aplicações multimodal. Em

suma, o SALT gerencia recursos de fala, reconhecimento de voz, sintetização da fala humana, DTMF, gravação de áudio e proporciona a flexibilidade para a reprodução de sons previamente gravados.

O formalismo e a sintaxe do SALT são independentes da natureza do documento fonte, sendo assim, ele pode ser usado efetivamente da mesma maneira dentro do HTML e em todas as suas variações, ou com WML e também com qualquer outra derivação de marcação SGML. Os principais elementos utilizados para o desenvolvimento de aplicações na linguagem SALT são o "*prompt*", o "*listen*", o DTMF e o "*smex*". O comando "*listen*" é composto de outros três elementos que são o "*grammar*", o "*bind*" e o "*record*".

Existem várias vantagens no que se refere, a utilização pelo SALT de uma linguagem madura como o HTML. Isto é bastante evidente, quando se trata dos eventos e dos modelos de roteiros, suportados pelos navegadores visuais que podem ser utilizados pelas aplicações SALT. Eles têm o objetivo de implementar fluxos de diálogos e outras formas de interação sem a necessidade de uma marcação extra, e possibilitando implementar a capacidade de fala em páginas visuais, permitindo de uma maneira simples e intuitiva a criação de aplicações multimodais. Desse modo, o SALT é uma ferramenta muito poderosa para a construção de interfaces de fala em aplicações que utilizam páginas da internet.

A linguagem SALT é especificada de acordo com os seguintes princípios fundamentais:

a) Integração transparente entre as interfaces de fala e páginas da internet.

O SALT é projetado para facilitar o uso de modelos DOM em páginas da internet. Para as aplicações multimodal ele integra de forma transparente as marcações das páginas visuais. Isso permite uma facilidade muito grande no processo de desenvolvimento de aplicações com o uso dessa linguagem.

b) Separação da interface de fala, da lógica do negócio e dos dados

Baseia-se no princípio da integração transparente, onde a interface de fala deve estar separada da lógica do negócio e dos dados. O SALT não pode estender diretamente qualquer linguagem de marcação individual, isto é, ele implementa uma interface de fala através de

uma camada separada da linguagem de marcação. Um diálogo SALT pode ser adaptado a uma estrutura de dados, por exemplo, um "form" HTML, então uma interface de fala ou um componente de um diálogo, pode ser reutilizado por outras páginas ou aplicações.

#### c) Potencialidade e flexibilidade do modelo de programação

A flexibilidade na programação das interfaces de fala é fundamental para que possam ser construídas aplicações de fala com qualidade. O SALT oferece um controle de alto nível na execução dos diálogos, através de eventos DOM e modelos de programas extremamente familiares aos desenvolvedores de aplicações. Isto permite ao Salt, com a utilização de seus elementos, fornecer aos projetistas um modelo simples e intuitivo.

#### d) Reuso de padrões existentes de gramática, saída de fala e semântica

O SALT reutiliza padrões já existentes de saída de fala, formatos de gramática e resultados semânticos, facilitando de forma significativa a construção de novas interfaces de fala.

#### e) Gama de dispositivos

O SALT não foi desenhado para um tipo de dispositivo específico, e sim para uma grande variedade de cenários. Dessa maneira implementa fala em páginas da internet de forma genérica, isto é, independente do dispositivo. Por exemplo, um computador pessoal pode executar localmente uma tarefa de reconhecimento de voz ou um processo de saída de fala, mesmo com uma capacidade limitada de processamento, devido ao SALT permitir a utilização de um servidor remoto para processar o reconhecimento de voz e serviços de sintetização da fala.

#### f). Custo mínimo de construção de programas por modo e por dispositivos

Derivado de todos os princípios anteriores, é uma noção que tem se tornado de fundamental importância para os desenvolvedores de página da internet, já que os seus clientes possuem vários tipos de dispositivos. Minimiza a sobrecarga na construção de programas para diferentes modos e diferentes clientes. Isso possibilita duas importantes classes de aplicações:

\* Multimodal - Quando uma página visual pode ser manipulada através de uma interface de fala em um mesmo dispositivo;

\* Cross-modal - Quando uma página de uma determinada aplicação pode ser reutilizada de diferentes modos em diferentes dispositivos. Por exemplo, utilização em uma interface apenas visual ou de fala.

#### 4.3.3 Comparação entre VoiceXML e SALT

A análise comparativa entre essas duas linguagens não é algo elementar, mesmo porque, irá depender de vários fatores que ao serem definidos e esclarecidos, contribuirão para a melhor tomada de decisão.

Perguntas como as descritas a seguir auxiliarão em muito esse processo, bem como as respectivas respostas para essas questões, são de fundamental importância no momento de decidir qual será a linguagem que mais se adapta as necessidades da organização, no que se refere a construção de aplicações de voz. Portanto, o objetivo principal não deve ser o de se encontrar a melhor, ou a pior linguagem, e sim, a mais adequada para a sua necessidade.

##### 4.3.3.1 Plataforma atual de desenvolvimento

Se toda a infra-estrutura de desenvolvimento de aplicativos da organização está baseada em .NET da Microsoft, a linguagem mais natural será o SALT, pois a plataforma de desenvolvimento SDK de fala da Microsoft se acopla perfeitamente com os componentes ASP e com os componentes .NET do Microsoft Visual Studio. Já as organizações que tem a sua infra-estrutura de desenvolvimento baseada em J2EE devem optar pelo VoiceXML.

Tecnicamente ambas as linguagens são compatíveis com qualquer servidor *WEB*. Porém, as ferramentas de desenvolvimento incluídas pelos fornecedores de soluções VoiceXML são normalmente baseadas no Java, enquanto as ferramentas incluídas para o SALT são obviamente o .NET. É possível adaptar as várias soluções para as duas linguagens de desenvolvimento, mas esse fator freqüentemente provoca aumento no tempo de execução e no custo total do projeto. Outro aspecto relevante, é que a aplicação desenvolvida em VoiceXML tem alguma vantagem sobre as desenvolvidas em SALT, devido a maturidade da linguagem

VoiceXml, que está a mais tempo no mercado.

#### 4.3.3.2 Fornecedores da Solução

Outro fator que deve ser levado em conta, no momento de definir a linguagem a ser utilizada, é o nível de interação dessas linguagens, com as soluções de voz propostas pelos fornecedores existentes atualmente no mercado. O VoiceXML oferece algumas vantagens no que se refere ao tempo para a execução do projeto e na diversidade dos fornecedores para a solução de automação de voz. Esse fato ocorre, principalmente, porque o VoiceXML está a mais tempo no mercado e a grande maioria dos provedores desse tipo de solução são compatíveis com seu padrão. O SALT, por estar menos maduro, tem uma menor gama de fornecedores integrados ao seu padrão.

As vantagens acima relatadas, não fecham a questão sobre a tendência do mercado para construção de aplicações de voz. Apesar de vários aspectos apontarem favoravelmente no momento, para o VoiceXML, algo que deve ser levado em conta, é o poder de investimento e formação de cultura, ou seja, alto grau de aceitação dos seus produtos, que o fabricante da linguagem SALT tem no mercado de desenvolvimento de aplicações.

#### 4.3.3.3 A aceitação dos profissionais

Ambas as linguagens, por estarem focadas em um tipo específico de produto, que é a construção de aplicações de voz, atingem apenas um pequeno grupo de profissionais. Portanto, se o objetivo for o de desenvolver aplicações com maior rapidez, o VoiceXML, mais uma vez sai na frente, devido a sua maturidade no mercado. A possibilidade de se encontrar um profissional com essas características é muito maior, do que um profissional que conheça SALT.

Existem também, algumas características individuais, de cada linguagem que facilitam o desenvolvimento de aplicativos. O VoiceXML tem funções de controle de fluxo, já o SALT depende de outras linguagens para controlar o seu fluxo como o HTML. As ferramentas que auxiliam o desenvolvimento em SALT são mais amigáveis, facilitam o seu aprendizado e

proporciona maior liberdade ao desenvolvedor. As aplicações quando desenvolvidas para telefonia, baseadas na fala humana, são mais simples ao serem desenvolvidas na linguagem VoiceXML. No entanto, nos casos de aplicações multimodal o SALT oferece mais recursos ao desenvolvedor.

#### 4.3.3.4 Característica da aplicação a ser desenvolvida

Inicialmente, cabe definir com maior clareza o conceito de aplicação multimodal. Uma aplicação somente é considerada com funções de multimodal, quando permite a utilização de vários dispositivos, como telefone, PDAs e outros, e disponibiliza diversas formas de interação de entrada e de saída das informações, como a voz e a imagem. Por exemplo, ao utilizar uma aplicação geográfica com funções multimodal, o usuário através da fala executaria comandos de direcionamento em um mapa que está sendo visualizado em um dispositivo de exibição qualquer, e a orientação de qual é esse local no mapa seria informada através de uma interface audível.

Tanto o VoiceXML quanto o SALT têm características multimodais. Porém, o SALT por contemplar essa característica desde a sua concepção permite a construção de aplicações multimodal de forma mais natural. Já o VoiceXML, que tem o seu ponto forte em aplicações de telefonia com fala, também permite a construção de aplicações multimodal, mas exige maior esforço durante o processo de desenvolvimento.

#### 4.4 Principais Plataformas de Desenvolvimento desse Tipo de Tecnologia no Brasil

A diversidade das soluções que implementam os padrões VoiceXML ou SALT em seus produtos e as várias arquiteturas atualmente existentes no mercado nacional dificulta essa abordagem, principalmente ao se levar em conta a complexidade da infra-estrutura necessária e a enorme gama de ferramentas que contemplam esse tipo de solução.

Porém, o foco será nos principais produtos e seus respectivos fornecedores dessa tecnologia na construção de portais de voz.

##### 4.4.1 Nuance V-Builder

A Nuance, que é também fornecedora de software de interface de voz, desempenha papel relevante na construção da infra-estrutura para portais de Voz, fornecendo diversas tecnologias-chave, tais como: interface de padrão de usuário, desenvolvimento de ferramentas de produtividade, autenticação de voz, arquitetura de rede escalável e evolução dos padrões.

No caso da interface padrão de usuário, a Nuance desenvolveu um navegador da internet que compreende comandos de voz. O produto foi projetado para padronizar a interface de usuário em qualquer aplicação comandada por voz e é similar em funcionalidade aos produtos atualmente existentes. Como exemplo de ferramenta de produtividade já desenvolvida, a Nuance apresenta hoje a Nuance *V-Builder*, que permite aos desenvolvedores criarem interativamente interfaces de voz para os sítios da internet, ou qualquer outro tipo de aplicação.

O Nuance Verifier é o '*software*' que permite ao usuário ser reconhecido e autenticado a partir de sua voz. Isso significa que a combinação do reconhecimento com a autenticação da voz assegura um nível extremamente elevado de segurança, que elimina a necessidade de senhas ou códigos de identificação.

#### 4.4.2 WebSphere Voice Server

O *WebSphere Voice Server* fornece a infra-estrutura para que aplicações sejam acessadas por comandos de voz, usando gerenciamento de diálogos simples, possibilitando que os usuários utilizem a fala como uma forma de interface. Os sistemas de reconhecimento de voz permitem rápida redução de custos em empresas que utilizem estruturas de CRM e dependem da interação constante com clientes, parceiros ou funcionários

Os recursos do WebSphere Voice Server podem ser usados para desenvolver e testar aplicações de voz e incluem suporte para aplicações escritas em Voice XML, JavaBeans ou C. O *WebSphere Voice Toolkit* permite que desenvolvedores acrescentem tecnologia de voz a diversas aplicações e ainda inclui um editor de Voice XML, editor de gramática e um editor de pronúncias. O *WebSphere Voice Server Software Developers Kit (SDK)*, fornece ferramentas para a construção de protótipos de aplicativos Voice XML em um computador pessoal, sem a necessidade de um servidor de telefonia.

#### 4.4.3 OracleAS 10G

A Oracle como provedora de soluções de armazenamento de dados e no desenvolvimento de aplicações, também fornece uma família de produtos que contemplam a interação entre o usuário final e o aplicativo, através do reconhecimento da voz humana. O Oracle tem alguns módulos que permitem ao desenvolvedor de sistemas, utilizar todas as vantagens e as facilidades que existem dentro desse produto para a criação de portais de voz.

As principais características e os princípios básicos de funcionamento do OracleAS 10G, no que se refere a aplicações de voz, são as seguintes:

Inicialmente, na utilização de uma aplicação via voz, existe a necessidade de que um usuário final faça uma solicitação para um provedor desse tipo de serviço, através de uma linha telefônica tradicional, e que haja uma interação com uma aplicação audível. O usuário final tem dois métodos possíveis de entrada que podem ser executados através da fala ou do teclado do telefone

Antes do processo atual de reconhecimento da fala, o usuário necessitava treinar o programa para reconhecer as suas expressões vocais. Hoje, esse processo está suficientemente maduro para que esse reconhecimento da fala humana seja mais transparente, não havendo a necessidade desse treinamento, e em conjunto com linguagens de programação como o Voice XML, possam ser largamente utilizados para a criação de sistemas complexos de reconhecimento de voz e no desenvolvimento de aplicações que usam o telefone como meio de interação. Existem três componentes básicos que fazem parte dessa arquitetura: o interpretador de voz, o servidor de aplicações, e a fonte de dados.

O servidor de aplicação é responsável pela geração das páginas VoiceXML para o OracleAS 10G. Essas páginas VoiceXML normalmente permitem ao usuário executar tarefas com a utilização de comandos de voz, e possibilitam ao usuário criar e armazenar em um banco de dados um perfil personalizado. No decorrer de uma ligação telefônica, o servidor de aplicação pode gerar várias páginas VoiceXML, porém essa geração de páginas é transparente para o usuário.

Para controlar uma determinada entrada de chamada de um usuário, o interpretador de voz



aciona páginas VoiceXML do servidor de aplicação que as interpreta. A página VoiceXML normalmente contém um conjunto de registros de áudio denominados '*prompts*' que permitem ao usuário fazer uma comunicação com o aplicativo através da fala. O interpretador VoiceXML utiliza um componente de reconhecimento de voz, como o da Nuance, para reconhecer os comandos de voz executados pelo usuário. O interpretador então pode efetivamente tomar as medidas necessárias para atender essas solicitações. Essas ações podem ser a de reproduzir o som de um arquivo de áudio, fazer uma pergunta ao usuário, ou a de acionar uma outra página VoiceXML. O OracleAS 10G não contém o componente de interpretação de voz; Entretanto ele funciona com qualquer interpretador compatível com o VoiceXML, incluindo o NMS, Verascape, Comverse, IBM, e Voice Genie.

#### 4.4.4 Microsoft .NET

A Microsoft sendo uma das principais empresas que desenvolveram o SALT, tem uma grande participação no processo de desenvolvimento de aplicações de voz, provendo ferramentas e soluções, através de uma enorme gama de produtos próprios, ou mesmo, através de parcerias com outras empresas. Como exemplo de uma parceria bem sucedida é a união da Microsoft com a Interoice, que juntas estão tornando a criação e o desenvolvimento de soluções automatizadas com o uso da voz humana, mais rápidas, fáceis e econômicas, ao serem utilizadas em larga escala. Entre elas, podem ser citadas aplicações de Interação com Resposta de Voz (IVR), Administração de Relacionamento com o Cliente (CRM) e Planejamento de Recursos Empresariais (ERP).

Como parte dessa aliança, a Interoice é a responsável por conduzir um programa de desenvolvimento junto com a Microsoft. Os desenvolvedores da Interoice e seus projetistas de interfaces para o usuário ganharam um grande aliado no processo de desenvolvimento das aplicações de voz ao utilizar o Microsoft .NET, como plataforma para construção de sistemas baseados no reconhecimento da fala humana.

Em resumo, a Microsoft fornece um conjunto de ferramentas que permitem aos desenvolvedores de sistemas baseados em comandos de voz, criarem aplicativos de forma simples e intuitiva, através do uso de uma família de produtos como o Omvia (componente de fluxo de chamada por fala), o ASP (Teste de Aplicação e Certificação), o pacote SALT (Desenvolvimento VUI, MSSAS SDK, Servidor de Fala), entre outros.

A seção forneceu uma visão geral do custo de um processo de construção de um portal de voz, em especial, os problemas encontrados durante o processo de migração e as suas possíveis soluções. Após essa definição, será necessário implementá-la. Portanto, foi apresentada de uma forma sucinta as duas principais linguagens para construção de aplicações de voz e as plataformas de desenvolvimento existentes no país. A próxima seção fará a implementação da MoLIC para o IPVA no portal de voz com foco em governo eletrônico.

## 5. IMPLEMENTAÇÃO DA MODELAGEM DO IPVA PARA O PORTAL DE VOZ GOVERNAMENTAL

As seções anteriores forneceram uma noção de como a tecnologia de portais de voz podem ser úteis na construção de aplicações de cunho governamental. Entretanto, o processo de conversão dos atuais sítios de governo eletrônico para essa nova tecnologia não é uma tarefa muito simples.

A concepção desses aplicativos está baseada em uma interface visual que permite conduzir o usuário através de imagens e textos ao seu objetivo final. Isso permite que o usuário navegue intuitivamente, entre as várias opções fornecidas nesses sítios. Já no portal de voz, o processo é diferente, a única forma de comunicação é o som, que deve fornecer opções objetivas e conter uma fraseologia muito semelhante a linguagem natural do usuário citado por BRANCO et al (2006)

Outra coisa que cabe ser ressaltada, é que a quantidade de informações em uma única página pode ser maior, pois não necessita da memorização das diversas opções existentes por parte do usuário final. Contudo, em um portal de voz não existe o aspecto visual, dessa forma, a árvore de opções deve ser menor, evitando uma lista muito grande de itens a serem escolhidos conforme cita Gruenstein(2002). Em suma, esses e outros fatores já abordados anteriormente, comprovaram que a migração dos sítios atuais para a tecnologia de portais de voz é algo muito complexo e necessita ser muito bem concebida, para não provocar a rejeição de uso por parte do usuário final.

A MoLIC conforme cita Paula(2003) sendo uma linguagem fundamentada em conceitos da engenharia semiótica, se transformou em um grande aliado na modelagem do portal de voz com foco em governo eletrônico. O seu foco principal é no processo de comunicação entre o aplicativo e o usuário, privilegiando os diálogos entre esses dois elementos.

### 5.1 Modelagem através da MoLIC para o IPVA

O portal de voz governamental tem uma enorme gama de funcionalidades e de serviços a serem oferecidos para a população em geral. Portanto, esses fatores fazem com que ele se torne um aplicativo grande e bastante complexo.

A concepção de uma modelagem para todo o portal de voz, seria algo impraticável dentro desse estudo. Sendo assim, a proposta é a de usar a MoLIC para modelar o nível de transação para uma aplicação típica, no caso, o pagamento do Imposto sobre Propriedade de Veículos Automotores (IPVA).

Esta transação é uma excelente candidata para o propósito em questão, pois tem todas as características de uma transação de governo eletrônico – autenticação de usuário, fornecimento de dados pessoais, emissão de comprovantes e prestação de informações – em um sistema com nível médio de complexidade. Como simplificação, esta transação não tem um alto grau de dependência com outros aplicativos. Outro fator favorável é que o pagamento do IPVA é um dos serviços mais acessados do Portal do Cidadão e um dos impostos estaduais mais utilizado e conhecido pelos cidadãos do estado de São Paulo.

A MoLIC, como toda linguagem de modelagem, está baseada em diagramas (figuras) para representar os seus elementos, fluxos e interações. Porém, esse tipo de representação é um fator impeditivo para as pessoas portadoras de deficiência visual parcial ou total, limitando de forma significativa a sua utilização.

Com o objetivo de tornar a MoLIC acessível para todos, baseado em uma experiência pessoal, a proposta desse estudo é a de substituir esses diagramas por uma estrutura hierárquica e textual. Cabe ressaltar, que todos os termos e nomenclaturas da MoLIC foram preservadas para manter compatibilidade com a linguagem original.

#### 5.1.1 Cenários Identificados para o IPVA

A proposta é a de criar um conjunto de cenários que identifiquem, senão todas, as principais interações existentes no sistema do IPVA. Eles foram construídos baseados nas simulações de uso identificadas como sendo as mais importantes funcionalidades desse aplicativo.

Os cenários, conforme cita Paula(2003), têm o objetivo de representar um conjunto de atividades executadas pelo usuário para atingir uma determinada meta. Eles descrevem um modelo de tarefas humanas que especifica os requisitos do usuário, sem definir quais serão os objetos utilizados nesse diálogo.

Contudo, no caso do portal de voz, o foco principal é o de atingir durante o processo de comunicação uma interação muito próxima da linguagem natural do ser humano. Portanto, é fundamental que o projetista fique muito atento para não implementar de forma inadequada esses conceitos, confundindo requisitos de interação com os objetos do diálogo.

Com o objetivo de facilitar o entendimento dos cenários relatados, segue uma breve descrição das principais atividades neles realizadas.

O cenário 1 narra a navegação de um usuário através do portal de voz governamental, focando a obtenção de informações sobre o IPVA. Nele são validadas funcionalidades como a não entrada de informações entre o usuário e o portal, a inexperiência no uso e todas as etapas para se obter informações sobre o IPVA.

O cenário 2 descreve todas as etapas para cadastrar um usuário no portal, pois a consulta do saldo do IPVA para um determinado veículo, necessita que o usuário esteja cadastrado. Nele são identificadas tarefas como a autenticação biométrica através da voz, o cadastramento de usuários, correção de dados cadastrais e optar pelo tipo de emissão do comprovante.

O cenário 3 faz a narrativa de um processo típico do pagamento do IPVA. Nele são validadas funcionalidades como o login, o valor a ser pago, a forma de pagamento e o tipo de emissão do comprovante.

O cenário 4 descreve o pagamento do IPVA para um veículo novo, através do portal de voz governamental. Nele são identificadas tarefas como o login, o cadastramento do veículo, o cadastramento do usuário proprietário, a obtenção do valor a ser pago, a forma de pagamento e o tipo de emissão do comprovante.

O cenário 5 narra o cadastramento de dúvidas sobre o IPVA, tornando o portal de voz governamental, algo maior do que um simples local para o cidadão executar os seus pagamentos (transações). Ele também proporciona ao usuário, a possibilidade do esclarecimento de dúvidas sobre os temas mais diversos, dentro da sua proposta principal, que é a de prestar um conjunto de serviços estaduais para a população em geral. Nele são identificadas tarefas como o atalho para se atingir o serviço de forma mais eficiente, o

cadastro do usuário, o cadastro da dúvida e o tipo de emissão do aviso do esclarecimento da dúvida.

O cenário 6 descreve uma alteração nos dados cadastrais do usuário e do proprietário. Cabe observar, que a atualização no dado cadastral do usuário do portal, foi automaticamente refletida para o cadastro de proprietários de veículos. Nele são verificadas tarefas como o login, a alteração dos dados cadastrais do usuário do portal e a alteração dos dados cadastrais do proprietário.

O cenário 7 aborda a utilização da entrada de vários atributos de pesquisa em um único diálogo, proporcionando um acesso muito mais rápido as informações desejadas. Nele são identificadas tarefas como o atalho e a consulta ao valor do IPVA.

O cenário 8 descreve o processo de solicitação de ajuda no portal de voz, que pode ser acionado a qualquer momento ou em qualquer local.

Nele são identificadas tarefas como repetir a última fala, voltar a interação anterior, sair do portal e falar com um atendente.

O cenário 9 descreve a consulta e a alteração dos dados cadastrais de um veículo dentro do portal de voz governamental. Nele são observadas tarefas como o login, a consulta aos dados do veículo, a alteração dos dados cadastrais do veículo e o tipo da emissão do comprovante.

O cenário 10 contém a narrativa de uma consulta aos dados cadastrais de um usuário dentro do portal de voz governamental. Nele são identificadas tarefas como o login e a consulta aos dados cadastrais do usuário.

O cenário 11 descreve a execução do pagamento dos anos anteriores do IPVA. Cabe observar, que em dado momento do diálogo, o usuário solicita diretamente ao portal para voltar a interação anterior e repetir a fala. Nele são verificadas tarefas como o login, consulta do valor a ser pago, a forma de pagamento e o tipo de emissão do comprovante.

Nos cenários que se seguem, as falas do portal de voz estão em itálico para facilitar a sua visualização. Outro aspecto que cabe ser ressaltado é a palavra “Diálogo” que representa o processo de comunicação do portal de voz, a ser utilizado na construção dos cenários.

Conforme recomendação da MoLIC, esta comunicação deve ser livre de contextualização e detalhes dos objetos utilizados no diálogo.

## Cenário 1- Informações sobre o IPVA

João ficou sabendo, através do seu amigo Arnaldo, da existência de um tal de portal de voz que fornece serviços governamentais para qualquer cidadão [1,4]. Ele gostaria de utilizá-lo, porém somente foi lhe dado um número de um telefone e ele não tem a mínima idéia de como proceder para obter determinado serviço.

“Bem, já que estamos no começo do ano, gostaria de sanar algumas dúvidas que tenho a respeito do IPVA.”

Depois de alguns dias João se encontrou com o Arnaldo e contou a sua experiência da seguinte forma:

João: Telefonei para aquele número que você me passou e os nossos diálogos foram os seguintes:

PVG: “*Diálogo: mensagem de boas vindas e optar pelo serviço desejado*” [2,3,5,7,8,29,34]

João: “Inicialmente fiquei sem saber o que fazer e não falei nada. Após alguns segundos ...” [3,5,6]

PVG: “*Diálogo: mensagem de boas vindas e optar pelo serviço desejado*” [2,3,5,7,8,29,34]

João: “Como eu não fiz nenhuma solicitação, pela segunda vez ...”

PVG: “*Diálogo: lista* contendo quais os tipos de serviços podem ser obtidos dentro do portal. “[3,5]

João: “Uma das opções de serviços foi a de "Impostos". Como eu sabia que o IPVA é um imposto estadual, falei "Impostos". Por incrível que pareça, o portal entendeu o que eu disse e voltou a perguntar:”

PVG: “*Diálogo: optar pelo tipo de imposto estadual*”

João: “Mais uma vez não soube o que fazer e esperei que o portal de voz me perguntasse pela segunda vez. Dado o meu silêncio, ele passou a dizer:”

PVG: “*Diálogo: lista* contendo os tipos de impostos estaduais”

João: “Gostei da coisa e me arrisquei novamente, e falei a sigla "IPVA". Heureka!!! O portal me entendeu novamente e passou a dizer:”

PVG: “*Diálogo: optar por efetuar pagamento, consulta, cadastro ou informações sobre as dúvidas mais comuns*”

João: “Falei “Informações sobre as dúvidas mais comuns” e passei a ouvir um texto com uma série de perguntas e respostas sobre o IPVA. Ótimo, lá estava o esclarecimento para as minhas dúvidas.”



## Cenário 2 - Cadastramento de um usuário

Isabella irá comprar um veículo usado e necessita saber se todos os IPVAs relativos aos anos anteriores foram pagos corretamente. Para tal, resolve acessar pela primeira vez o portal de voz governamental.

PVG: "*Diálogo: mensagem de boas vindas e optar pelo serviço desejado*" [2,3,5,7,8,29,34]

Isabella: "IPVA"

PVG: "Diálogo: optar por efetuar pagamento, consulta, cadastro ou informações sobre as dúvidas mais comuns"

Isabella: "consulta"

PVG: "Diálogo: optar por informações sobre o valor do IPVA ou o saldo devedor do veículo"

Isabella: "saldo devedor"

PVG: "Diálogo: informar o nome completo"[11]

Isabella: "Isabella Carvalho"

PVG: "Erro: usuário não encontrado"

"Diálogo: optar por tentar novamente ou adicionar um novo usuário" [9,13]

Isabella: "Adicionar usuário"

PVG: "Diálogo: informar o nome completo"

Isabella: "Isabella Carvalho"

PVG: "Diálogo: informar o endereço completo"

Isabella: "Rua das Andorinhas, 30"

PVG: "Diálogo: digitar o número do CEP"

"Diálogo: digitar o número do CPF ou do CNPJ"

"Diálogo: informar o e-mail"

Isabella: "icarvalho@ig.com.br"

PVG: "Diálogo: digitar o número do FAX"

PVG: "Diálogo: Listar os dados fornecidos anteriormente"

"Diálogo: Optar por ouvir novamente (sim/não)"

Isabella: "não"

PVG: "Diálogo: confirmar Os dados cadastrais (sim/não)"

Isabella: "não"

PVG: "Diálogo: informar o dado errado"

Isabella: "O e-mail"

PVG: "Diálogo: informar o e-mail"

Isabella: "icarvalho@ig.com.br"

PVG: "Diálogo: listar O e-mail e solicitar confirmação"

Isabella: "sim"

PVG: "Diálogo: Optar por listar novamente os dados cadastrais (sim/não)"

Isabella: "não"

PVG: "Diálogo: informar a existência de mais algum dado errado (sim/não)"

Isabella: "não"

PVG: "Diálogo: sucesso no cadastro"

"Diálogo: optar por consultar o ano corrente ou os anos anteriores" [21,22]

Isabella: "os anos anteriores"

PVG: "Diálogo: digitar o código RENAVAL"

"Diálogo: listar O saldo devedor dos anos anteriores"

PVG: "Diálogo: optar pela emissão do comprovante por e-mail, fax ou correio" [12,25]

Isabella: "e-mail"

PVG: "Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"

### Cenário 3 - Pagamento do IPVA

Mariana gostaria de pagar o IPVA do seu carro com desconto e lembrou que hoje é o último dia para fazê-lo. Entretanto, os bancos já fecharam e ela está com problemas com o seu computador pessoal. E agora??? Parece que existe um portal de voz que oferece serviços governamentais, será que ele poderá me ajudar?

Através do guia de assinantes da telefonica (serviço 102) ela consegue obter o número do telefone do portal de voz. Posteriormente, ela liga para o portal e o diálogo ocorre da seguinte forma:

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]  
 Mariana: "Pagar o IPVA"  
 PVG: *"Diálogo: informar o nome completo"*[11]  
 Mariana: "Mariana Alves"  
 PVG: "Erro: usuário não encontrado"  
 "Diálogo: optar por tentar novamente ou adicionar um novo usuário" [9,13]  
 Mariana: "Adicionar usuário"  
 PVG: *"Diálogo: informar o nome completo"*  
 " ... "  
 PVG: *"Diálogo: sucesso no cadastro"*  
 PVG: *"Diálogo: Optar pelo pagamento deste ano ou dos anos anteriores"* [23,24]  
 Mariana: "deste ano"  
 PVG: *"Diálogo: optar pelo pagamento único ou parcelado"* [18]  
 Mariana: "pagamento único"  
 PVG: *"Diálogo: mensagem que informa o período de validade do desconto"* [30]  
 "Diálogo: digitar o código RENAVAL" [30]  
 PVG: *"Diálogo: mensagem com o valor a ser pago"*  
 "Diálogo: optar pela emissão de um boleto bancário ou pagar agora?" [19,20]  
 Mariana: "Hummm! ... pagar agora"  
 PVG: *"Diálogo: informar o número do banco e da agência"*  
 Mariana: "151 09369"  
 PVG: *"Erro: repetir a informação"*  
 Mariana: "Banco 151 e agência 09369"  
 PVG: *"Diálogo: digitar o número da conta"*  
 PVG: *"Diálogo: digitar a senha"*  
 PVG: *"Diálogo: mensagem com o valor a ser debitado (sim/não)"*  
 Mariana: "sim"  
 PVG: "Diálogo: optar pela emissão do comprovante por e-mail, fax ou correio" [12,25]  
 Mariana: "correio"  
 PVG: "Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"

#### Cenário 4 - IPVA para veículo novo

Isabella pensou melhor e resolveu comprar um veículo novo ao invés de um usado. Já que ela está cadastrada no portal de voz, optou por executar as tarefas de pagamento do IPVA através desse instrumento facilitador. Portanto, os diálogos foram os seguintes:

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]  
 Isabella: "IPVA"  
 PVG: *"Diálogo: optar por efetuar pagamento, consulta, cadastro ou informações sobre as dívidas mais comuns"*  
 Isabella: "efetuar pagamento"  
 PVG: *"Diálogo: informar o nome completo"*[11]  
 Isabella: "Isabella Carvalho"  
 PVG: *"Diálogo: optar pelo pagamento deste ano ou dos anos anteriores"* [23,24]  
 Isabella: "deste ano"  
 PVG: *"Diálogo: optar pelo pagamento único ou parcelado"* [18]  
 Isabella: "pagamento parcelado"  
 PVG: *"Diálogo: optar pela parcela (primeira, segunda ou terceira)"*  
 Isabella: "primeira"  
 PVG: *"Diálogo: digitar o código RENAVAM"*  
 PVG: *"Erro: mensagem contendo o código RENAVAM não encontrado"*  
*"Diálogo: Optar por tentar novamente ou adicionar um veículo"* [14]  
 Isabella: "cadastrar um veículo novo"  
 PVG: *"Diálogo: digitar o Código RENAVAM"*  
 PVG: *"Diálogo: informar a placa do veículo"*  
 Isabella: "ABC 1234"  
 PVG: *"Diálogo: digitar o ano de fabricação"*  
 PVG: *"Diálogo: informar a cor do veículo"*  
 Isabella: "verde"  
 PVG: *"Diálogo: informar o chassis do veículo"*  
 Isabella: "abc123de456f"  
 PVG: *"Diálogo: informar a marca do veículo"*  
 Isabella: "fiat"  
 PVG: *"Diálogo: informar o modelo do veículo"*  
 Isabella: "siena hlx"  
 PVG: *"Diálogo: informar o tipo de combustível do veículo"*  
 Isabella: "alcool"  
 PVG: *"Diálogo: digitar o número do CPF ou do CNPJ do proprietário"*  
 PVG: *"Diálogo: Listar os dados fornecidos anteriormente"*  
 "Diálogo: Optar por ouvir novamente (sim/não)"  
 Isabella: "não"  
 PVG: *"Diálogo: confirmar Os dados cadastrais (sim/não)"*  
 Isabella: "sim"  
 PVG: *"Diálogo: mensagem com o CPF ou CNPJ não encontrado no cadastro de proprietários"*  
*"Diálogo: optar por alterar o dado ou cadastrar um novo proprietário"* [16]  
 Isabella: "cadastrar o proprietário"  
 PVG: *"Diálogo: lista dos dados do usuário que serão transferidos para o cadastro de proprietários"*  
 "Diálogo: Optar por ouvir novamente (sim/não)"

Isabella: "não"

PVG: "Diálogo: confirmar Os dados cadastrais (sim/não)"

Isabella: "sim"

PVG: "*Diálogo: sucesso no cadastro do proprietário*"

*"Diálogo: sucesso no cadastramento do veículo"*

"Diálogo: mensagem com o valor a ser pago"

"Diálogo: optar pela emissão de um boleto bancário ou pagar agora?" [19,20]

Isabella: "pagar agora"

PVG: "*Diálogo: Informar o número do banco e da agência*"

Isabella: "Banco 151 e agência 09369"

PVG: "*Diálogo: digitar o número da conta*"

PVG: "*Diálogo: digitar a senha*"

PVG: "*Diálogo: mensagem com o valor a ser debitado (sim/não)*"

Isabella: "sim"

PVG: "Diálogo: optar pela emissão do comprovante por e-mail, fax ou correio" [12,25]

Isabella: "e-mail"

PVG: "Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"

## Cenário 5 - Atalho e cadastramento de dúvidas

João gostou muito do serviço de informações do portal de voz. Agora ele tem uma outra dúvida sobre o IPVA, mas não quer passar por todas as etapas do processo. Será que não consigo obter essa informação de uma forma mais direta? Bem ... irei tentar, pegou o telefone e ...

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

João: "IPVA dúvidas mais comuns"

PVG: *"Diálogo: informar a dúvida"* [26]

João: "Posso pagar a segunda parcela do IPVA sem ter pago a primeira?" [32]

PVG: *"Diálogo: resposta ..."*

"Diálogo: informar se existe outra dúvida"

João: "Existe a possibilidade de alterar os dados cadastrais de um veículo ou de um proprietário?" [15,17]

PVG: *"Diálogo: resposta ..."*

"Diálogo: informar se existe outra dúvida"

João: "Existe alguma data limite para optar pelo pagamento parcelado?" [31]

PVG: *"Diálogo: resposta ..."*

"Diálogo: informar se existe outra dúvida"

João: "Existe alguma forma de isenção do IPVA?"

PVG: *"Erro: optar por repetir a dúvida"*

João: "Existe alguma forma de isenção do IPVA?"

PVG: *"Erro: pergunta não encontrada. Cadastrar (sim/não)"* [27]

João: "sim"

PVG: *"Diálogo: informar o nome completo"*[11]

João: "João de Almeida"

PVG: "Erro: usuário não encontrado"

"Diálogo: optar por tentar novamente ou adicionar um novo usuário" [9,13]

João: "Adicionar usuário"

PVG: *"Diálogo: informar o nome completo"*

" ... "

PVG: *"Diálogo: sucesso no cadastro"*

PVG: "Diálogo: optar pela emissão do comprovante por e-mail, fax ou correio" [12,28]

João: "e-mail"

PVG: "Diálogo: sucesso na operação. Opstar por sair ou por acessar outro serviço do portal"

## Cenário 6 - Alteração dos dados cadastrais do Usuário/Proprietário

Mariana mudou de residência e gostaria de alterar os seus dados cadastrais dentro do portal de voz e executou as seguintes etapas:

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Mariana: "IPVA"

PVG: *"Diálogo: optar por efetuar pagamento, consulta, cadastro ou informações sobre as dívidas mais comuns?"*

Mariana: "cadastro"

PVG: *"Diálogo: optar pelo Cadastro de usuário, de proprietário ou de veículo?"*

Mariana: "usuário"

PVG: *"Diálogo: optar por Consultar, Adicionar ou alterar dados cadastrais do usuário?"* [9,10]

Mariana: "alterar"

PVG: *"Diálogo: informar o nome completo"*[11]

Mariana: "Mariana Alves"

PVG: *"Diálogo: Listar os dados fornecidos anteriormente"*

*"Diálogo: Optar por ouvir novamente (sim/não)"*

Mariana: "não"

PVG: *"Diálogo: informar o dado a ser alterado"*

Mariana: "endereço"

PVG: *"Diálogo: informar o endereço completo"*

Mariana: "Rua das Gaivotas, 150"

PVG: *"Diálogo: listar e solicitar confirmação do novo endereço"*

Mariana: "sim"

PVG: *"Diálogo: digitar o número do CEP"*

PVG: *"Diálogo: listar e solicitar confirmação do novo CEP"*

Mariana: "sim"

PVG: *"Diálogo: optar por mais alguma alteração cadastral"*

Mariana: "não"

PVG: *"Diálogo: confirmar a mesma alteração no cadastro de proprietário (sim/não)"* [17]

Mariana: "sim"

PVG: *"Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"*

### Cenário 7 - Atalho para os serviços do portal de voz

Arnaldo conversou com o seu amigo João e soube através dele, que poderia acessar informações sobre o IPVA de forma direta. Será que se eu passar os dados do meu veículo o portal de voz me entenderá? Bem, vou tentar ...

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Arnaldo: "Valor do IPVA para um Corsa sle ano 2004" [21,22]

PVG: *"Diálogo: informar o tipo de combustível do veículo"*

Arnaldo: "gasolina"

"Diálogo: optar por consultar o ano corrente ou os anos anteriores" [21,22]

Arnaldo: "o ano corrente"

PVG: *"Diálogo: lista contendo Os valores do IPVA"* [18,30,31]

"Diálogo: Optar por ouvir novamente (sim/não)"

Arnaldo: "não"

PVG: "Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"

## Cenário 8 - Solicitação de ajuda

Ana ficou sabendo através dos jornais que existe um portal de voz governamental, que presta serviços a população de uma maneira muito simples. Parece ser algo muito interessante, pensou ela, acho que irei conhecer e ....

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Ana: "ajuda" [34]

PVG: *"Diálogo: mensagem contendo explicações como segue:"*

*"se não souber como responder a uma determinada pergunta, aguarde, pois após a segunda vez, sem resposta, o portal irá ditar um conjunto de opções"*

*"Se desejar voltar a uma etapa anterior, diga VOLTAR"* [33]

*"Se desejar que o portal repita algo, diga REPETIR"*[35]

*"Se desejar sair do portal, diga SAIR"* [29]

*"Se desejar falar com dos nossos atendentes, diga FALAR COM O ATENDENTE"*[7]

Ana: "voltar"

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Ana: ...

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Ana: ...

PVG: *"Diálogo: lista contendo os serviços disponíveis"*

Ana: "sair"

PVG: *"Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"*



## Cenário 9 - Alteração dos dados cadastrais de um Veículo

Isabella recebeu por e-mail o comprovante do cadastramento do seu veículo que foi executado através do portal do governo eletrônico do estado de São Paulo. Ao comparar esses dados com os do documento oficial observou que a cor do veículo está incorreta. O que fazer ... Bem, deve ser possível alterar esses dados através do portal e ...

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Isabella: "IPVA"

PVG: *"Diálogo: optar por efetuar pagamento, consulta, cadastro ou informações sobre as dívidas mais comuns"*

Isabella: "cadastro"

PVG: *"Diálogo: optar por Cadastro de usuário, de proprietário ou de veículo"*

Isabella: "veículo"

PVG: *"Diálogo: optar por Consultar, Adicionar ou alterar dados cadastrais do veículo"*[14,15,36]

Isabella: "consultar"

PVG: *"Diálogo: informar o nome completo"*[11]

Isabella: "Isabella Carvalho"

PVG: *"Diálogo: digitar o código RENAVAL"*

*"Diálogo: lista contendo os dados do veículo:"*

*"Diálogo: Optar por ouvir novamente (sim/não)"*

Isabella: "não"

PVG: *"Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"*

Isabella: "alterar dados do veículo"

PVG: *"Diálogo: informar o dado a ser alterado"*

Isabella: "cor"

PVG: *"Diálogo: informar a cor do veículo"*

Isabella: "verde esmeralda"

PVG: *"Diálogo: listar e solicitar confirmação da nova cor do veículo"*

Isabella: "sim"

PVG: *"Diálogo: optar por mais alguma alteração cadastral"*

Isabella: "não"

PVG: *"Diálogo: optar pela emissão do comprovante por e-mail, fax ou correio"* [12,25]

Isabella: "correio"

PVG: *"Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"*

## Cenário 10 - Consulta aos dados cadastrais do Proprietário

Mariana alterou os dados cadastrais do seu usuário no portal, mas tem dúvidas quanto aos seus dados armazenados no cadastro de proprietário. Com o objetivo de sanar essas dúvidas entrou no portal de voz e ...

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Mariana: "cadastro de proprietário"

PVG: *"Diálogo: optar por Consultar, Adicionar ou alterar dados cadastrais do proprietário"* [36]

Mariana: "consultar"

PVG: *"Diálogo: informar o nome completo"*[11]

Mariana: "Mariana Alves"

PVG: *"Diálogo: Lista contendo os dados cadastrais"*

"Diálogo: Optar por ouvir novamente (sim/não)"

Mariana: "não"

PVG: "Diálogo: sucesso na operação. Optar por sair ou por acessar outro serviço do portal"

## Cenário 11 - Pagamento dos anos anteriores do IPVA

Mariana comprou um outro veículo e soube posteriormente que ele está com algumas parcelas do IPVA atrasadas. Preocupada com o fato, gostaria de obter o valor a ser pago o mais rápido possível. Bem, irei acessar o portal de voz do governo eletrônico e dependendo do valor pagarei o IPVA e depois cobrarei do antigo proprietário. Sendo assim, as etapas foram as seguintes:

PVG: *"Diálogo: mensagem de boas vindas e optar pelo serviço desejado"* [2,3,5,7,8,29,34]

Mariana: "Pagar o IPVA"

PVG: *"Diálogo: informar o nome completo"*[11]

Mariana: "Mariana Alves"

PVG: *"Diálogo: optar pelo pagamento deste ano ou dos anos anteriores"* [23,24]

Mariana: "dos anos anteriores"

PVG: *"Diálogo: optar pelo pagamento único ou parcelado"* [18]

Mariana: "pagamento parcelado"

PVG: *"Diálogo: digitar o código RENAVAL"*

PVG: *"Diálogo: lista contendo os valores a serem pagos"*

Mariana: "repetir"[35]

PVG: *"Diálogo: optar pela emissão de um boleto bancário ou pagar agora"* [19,20]

Mariana: "pagar agora"

PVG: *"Diálogo: informar o número do banco e da agência"*

Mariana: "voltar" [33]

PVG: *"Diálogo: optar pela emissão de um boleto bancário ou pagar agora"* [19,20]

Mariana: "boleto bancário"

PVG: *"Diálogo: optar pela emissão do comprovante por e-mail, fax ou correio"* [12,25]

Mariana: "correio"

PVG: *"Diálogo: sucesso na operação. Opção por sair ou por acessar outro serviço do portal"*

### 5.1.1.1 Perguntas Exploradas nos Cenários

As perguntas auxiliam no processo de definição dos objetivos e das metas a serem atingidas durante os diálogos existentes entre o aplicativo e o usuário. Elas permitem identificar interações importantes e servem como um elemento coadjuvante na criação dos cenários.

1. Quem irá utilizar o serviço do PVG?
2. Quais serão os serviços fornecidos pelo PVG?
3. O Que fazer para obter o serviço desejado?
4. Que tipo de informação é fornecida no PVG?
5. Quais são as formas de encontrar um determinado serviço no PVG?
6. O que fazer caso o usuário não entre com o serviço desejado?
7. Quais as situações que o usuário falará com um atendente?
8. Como estará organizado o PVG?
9. Como se cadastrar no PVG?
10. Como alterar os dados cadastrais do usuário?
11. Existe algum serviço que o usuário acessa somente após ser cadastrado?
12. O Retorno das solicitações feitas pelo usuário serão sempre por voz?
13. Como será feita a identificação do usuário?
14. Como cadastrar um veículo novo?
15. Haverá a possibilidade de alterar os dados cadastrais de um veículo?
16. Como cadastrar um proprietário?
17. Haverá a possibilidade de alterar os dados cadastrais de um proprietário?
18. Quais as formas de pagamento do IPVA?
19. Haverá a possibilidade de pagar oipva online?
20. Haverá a possibilidade de emitir um boleto para pagamento bancário?
21. Como obter informações sobre o IPVA do ano corrente?
22. Como obter informações sobre o IPVA de anos anteriores?
23. Haverá a possibilidade de pagar o IPVA do ano corrente?
24. Haverá a possibilidade de pagar o IPVA de anos anteriores?
25. Quais serão as formas de emissão dos comprovantes?
26. Quais serão as formas de sanar dúvidas?
27. Haverá a possibilidade de cadastrar uma dúvida?
28. Como saber se a dúvida já foi respondida?
29. Existe a possibilidade de sair a qualquer momento?
30. Existe uma data limite para o pagamento com desconto?
31. Existe uma data limite para optar por pagamento parcelado?
32. É possível pagar a parcela atual sem ter pago a anterior?
33. Existe uma forma de retornar a etapa anterior?
34. Existe a possibilidade de solicitar ajuda?
35. Existe a possibilidade de voltar a ouvir uma lista de opções?
36. Existe a possibilidade de consultar dados cadastrais?

### 5.1.1.2 Tabela de Signos

Os signos são elementos que pertencem a aplicação e foram identificados durante a criação dos cenários para o sistema do IPVA.

#### Signo Composto: Usuário

nome	nome/identificador do usuário	transformado
Email	email do usuário	domínio
endereço	endereço do usuário	domínio
cep	número do cep do usuário	domínio
telefone	telefone do usuário	domínio
fax	fax do usuário	Domínio
cpf/cnpj	identificação do usuário	Domínio

#### Signo Composto: Veículo

RENAVAM	código RENAVAM	aplicação
placa	número da placa do veículo	domínio
ano de fabricação	data de fabricação do veículo	domínio
cor	cor do veículo	domínio
tipo	tipo do veículo	aplicação
chassis	número do chassis do veículo	domínio
marca	marca do veículo	domínio
modelo	modelo do veículo	domínio
combustível	tipo de combustível do veículo	domínio
cpf/cnpj	identificação do proprietário	aplicação

#### Signo Composto: Proprietário

nome	nome do proprietário	domínio
endereço	endereço do proprietário	domínio
cep	número do cep do proprietário	domínio
cpf/cnpj	identificação do proprietário	domínio

## Signo Composto: IPVA

ano	ano do ipva	domínio
data_1	data da primeira parcela	domínio
valor_1	valor da primeira parcela	domínio
data_2	data da segunda parcela	domínio
valor_2	valor da segunda parcela	domínio
data_3	data da terceira parcela	domínio
valor_3	valor da terceira parcela	domínio
RENAVAM	código do RENAVAM do veículo	aplicação
valor	valor do ipva	domínio
taxa desc	taxa de desconto	domínio
taxa multa	taxa da multa	domínio
saldo	saldo débito/crédito do ipva	aplicação

## Signo Composto: Tabela de Marca/Modelo

ano	ano do ipva	domínio
marca	marca do veículo	domínio
modelo	modelo do veículo	domínio
ano_fabricação	ano de fabricação	domínio
tipo	tipo de combustível do veículo	domínio
valor	valor venal do veículo	domínio
valor_ipva	valor do IPVA do veículo	domínio

## Signo Composto: Banco

banco	código do banco	domínio
agência	código da agência	domínio
conta	número da conta	domínio
valor	valor do débito	aplicação

## Signo Composto: Dúvidas

código	código da pergunta/resposta	aplicação
pergunta	Texto da pergunta	domínio
resposta	texto da resposta	domínio
data	data da dúvida	aplicação
status	status da pergunta (respondida ou não)	aplicação
cpf/cnpj	identificação do usuário	aplicação

### Signo Composto: Árvore de Serviços

serviço	tipo de serviço existente no portal	domínio
opção	Opções existentes para obter um determinado serviço	Aplicação
atalho	Atalhos existentes para obter diretamente um determinado serviço	Aplicação

#### 5.1.2 Modelo de Tarefas do IPVA

O modelo de tarefas consiste em extrair dos cenários as metas e os objetivos que o usuário tem ao utilizar o aplicativo. No portal de voz governamental, as metas identificadas no sistema do IPVA foram obtidas baseadas nos cenários.

##### 5.1.2.1 Metas do Usuário

As metas são compostas por alguns elementos que as definem e fornecem informações para que o projetista possa conceber o modelo de tarefas. Esses elementos são os critérios de classificação das metas, os papéis e as situações de uso identificadas através dos cenários, e dependem do tipo da aplicação. No caso do IPVA elas são:

Critério de Classificação das Metas: Utilização do portal de voz e Suporte ao portal de voz

Papéis:

U - Usuário não cadastrado

UC - Usuário cadastrado

UP - Usuário Proprietário

Situações de uso identificadas no portal de voz para o IPVA:

1. Optar pelo Serviço Desejado - Meta Optar pelo Serviço
2. Consultar IPVA - Árvore Utilização do Portal
3. Pagar IPVA - Árvore Utilização do Portal
4. Manipular cadastro -Árvore Suporte ao Portal

## 5. Esclarecer Dúvidas - Árvore Suporte ao Portal

### DIAGRAMA HIERÁRQUICO DE METAS

Portal de Voz Governamental

entrar no Portal

Optar pelo Serviço - A

#### 1. Utilizar o portal

##### 1.1. consultar IPVA

1.1.1 Valor do IPVA - B

1.1.2 Saldo Devedor - C

##### 1.2. pagar IPVA

1.2.1 Ano Corrente - D

1.2.2 Anos Anteriores - E

#### 2. Suporte ao Portal

##### 2.1. manipular Cadastro

##### 2.1.1 Usuário do Portal

2.1.1.1 Consultar - F

2.1.1.2 Adicionar - G

2.1.1.3 Alterar - h

##### 2.1.2 Proprietário

2.1.2.1 Consultar- I

2.1.2.2 Adicionar - J

2.1.2.3 Alterar - K

##### 2.1.3 Veículo

2.1.3.1 Consultar - L

2.1.3.2 Adicionar - M

2.1.3.3 Alterar - N

##### 2.2. Esclarecer Dúvidas

2.2.1 Acessar Dúvidas - O

2.2.2 Cadastrar Dúvidas - P

Modelo Diagramático de Tarefas:

Meta (A) - Optar pelo Serviço

1. Navegar pela Árvore de Opções

2. Acessar o Serviço através do Atalho

Meta (B) - Consultar Valor do IPVA

1. Fornecer Critérios de Pesquisa

2. Obter Resultado

2.a. Ano Corrente

2.b. Anos Anteriores

Meta (C) - Consultar Saldo Devedor do IPVA

Login

1. Fornecer Identificador Único do Veículo

2. Obter Resultado



- 2.a. Ano Corrente
- 2.b. Anos Anteriores
- 3. Emitir Comprovante
- 3.a. Forma de envio

Meta (D) - Pagar Ano Corrente do IPVA

Login

- 1. Fornecer Identificador Único do Veículo
- 1.a. Tipo de Pagamento
  - 1.a.1. Único com desconto
  - 1.a.2. Único sem desconto
  - 1.a.3. Parcelado
- 2. Forma de pagamento
  - 2.a. Imediato
  - 2.b. Posterior
- 3. Emitir Comprovante
- 3.a. Forma de envio

Meta (E) - Pagar Anos Anteriores do IPVA

Login

- 1. Fornecer Identificador Único do Veículo
- 1.a. Tipo de Pagamento
  - 1.a.1. Único
  - 1.a.2. Parcelado
- 2. Forma de pagamento
  - 2.a. Imediato
  - 2.b. Posterior
- 3. Emitir Comprovante
- 3.a. Forma de envio

Meta (F) - Consultar Usuário do Portal

Login

- 1. Informar Usuário
- 2. Exibir Dados do Usuário

Meta (G) - Alterar Usuário do Portal

Login

- 1. Informar Usuário
- 2. Exibir dados do Usuário
- 3. Fornecer novos dados do usuário

Meta (H) - Adicionar Usuário do Portal

- 1. Informar dados do usuário

Meta (I) - Consultar Proprietário

Login

- 1. Informar Proprietário
- 2. Exibir Dados do Proprietário

Meta (J) - Alterar Proprietário

## Login

1. Informar Proprietário
2. Exibir dados do Proprietário
3. Fornecer novos dados do Proprietário

## Meta (K) - Adicionar Proprietário

## Login

1. Informar dados do Proprietário

## Meta (L) - Consultar Veículo

## Login

1. Informar Veículo
2. Exibir Dados do Veículo

## Meta (M) - Alterar Veículo

## Login

1. Informar Veículo
2. Exibir dados do Veículo
3. Fornecer novos dados do Veículo
4. Emitir Comprovante
  - 4.a. Forma de envio

## Meta (N) - Adicionar Veículo

## Login

1. Informar dados do Veículo
2. Emitir Comprovante
  - 2.a. Forma de envio

## Meta (O) - Esclarecer Dúvida

1. Fornecer critério de pesquisa
2. Exibir resultado

## Meta (P) - Cadastrar Dúvida

## Login

1. Informar dúvida
  2. Emitir comprovante
    - 2.a. Forma de envio
- 5.1.2.2 Modelo Diagramático de Tarefas

O modelo Diagramático de tarefas consiste em uma decomposição hierárquica das metas do usuário. Ele descreve a estrutura das tarefas que serão executadas no aplicativo.

## Meta A: Optar pelo serviço

Cenários Associados: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

Papéis: U, UC e UP

A.1. Navegar pela árvore de opções

SIGNOS: árvore de serviços.opção?

A.2. Acessar o serviço através do atalho

SIGNOS: árvore de serviços.atalho?

A.3. Examinar Resultado

SIGNOS: Conjunto(árvore de serviços.serviço, 0..n)!  
 TRATAMENTO APOIADO: falta de informações obrigatórias

Meta B: Consultar valor do IPVA

Cenários Associados: 7

Papéis: U, UC e UP

B.1. Fornecer Critérios de Pesquisa

SIGNOS: tabela de marca/modelo.ano\_fabricação?, tabela de marca/modelo.marca?, tabela de marca/modelo.modelo?,

tabela de marca/modelo.tipo?, tabela de marca/modelo.ano?

TRATAMENTO APOIADO: falta de informações obrigatórias

B.2. Obter Resultado

B.2.a. Ano Corrente

SIGNOS: Conjunto(tabela de marca/modelo.valor\_ipva, 0..1)!

B.2.b. Anos Anteriores

SIGNOS: tabela de marca/modelo.valor\_ipva, 0..n!

Meta C: Consultar saldo devedor do IPVA

Cenários Associados: 2

Papéis: UC e UP

C.1. Fornecer Identificador Único do Veículo

SIGNOS: veículo.RENAVAM?

PREVENÇÃO ATIVA: somente usuários cadastrados ou proprietários de veículo podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

C.2. Obter Resultado

C.2.a. Ano Corrente

SIGNOS: Conjunto(ipva.\*, 0..1)!

C.2.b. Anos Anteriores

SIGNOS: Conjunto(ipva.\*, 0..n)!

C.3. Emitir Comprovante

C.3.a.1. Por e-mail

SIGNOS: usuario.e-mail?

C.3.a.2. Por fax

SIGNOS: usuario.fax?

C.3.a.3. Por correio

SIGNOS: usuario.endereço?, usuario.cep?

Meta D: Pagar ano corrente do IPVA

Cenários Associados: 3, 4

Papéis: UC e UP

D.1. Fornecer Identificador Único do Veículo

SIGNOS: veículo.RENAVAM?

PREVENÇÃO ATIVA: somente usuários cadastrados ou proprietários de veículo podem executar esse tipo de consulta

TRATAMENTO APOIADO: falta de informações obrigatórias

D.1.a. Tipo de Pagamento

D.1.a.1. Único com desconto

SIGNOS: ipva.valor, 0..1!

D.1.a.2. Único sem desconto

SIGNOS: ipva.valor, 0..1!

D.1.a.3. Parcelado

SIGNOS: conjunto(ipva.data\_1, ipva.valor\_1, ipva.data\_2, ipva.valor\_2, ipva.data\_3, ipva.valor\_3, 0..1)!

D.2. Forma de pagamento

D.2.a. Imediato

SIGNOS: banco.banco?, banco.agência?, banco.conta?, banco.senha?

TRATAMENTO APOIADO: falta de informações obrigatórias

D.2.b. Posterior

D.3. Emitir Comprovante

D.3.a.1. Por e-mail

SIGNOS: usuario.e-mail?

D.3.a.2. Por fax

SIGNOS: usuario.fax?

D.3.a.3. Por correio

SIGNOS: usuario.endereço?, usuario.cep?

Meta E: Pagar anos anteriores do IPVA

Cenários Associados: 11

Papéis: UC e UP

E.1. Fornecer Identificador Único do Veículo

SIGNOS: veículo.RENAVAM?

PREVENÇÃO ATIVA: somente usuários cadastrados ou proprietários de veículo podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

E.1.a. Tipo de Pagamento

E.1.a.1. Único

SIGNOS: ipva.valor, 0..1!

E.1.a.2. Parcelado

SIGNOS: conjunto(ipva.data\_1, ipva.valor\_1, ipva.data\_2, ipva.valor\_2, ipva.data\_3, ipva.valor\_3, 0..1)!

E.2. Forma de pagamento

E.2.a. Imediato

SIGNOS: banco.banco?, banco.agência?, banco.conta?, banco.senha?

TRATAMENTO APOIADO: falta de informações obrigatórias

E.2.b. Posterior

E.3. Emitir Comprovante

E.3.a.1. Por e-mail

SIGNOS: usuario.e-mail?

E.3.a.2. Por fax

SIGNOS: usuario.fax?

E.3.a.3. Por correio

SIGNOS: usuario.endereço?, usuario.cep?

Meta F: Consultar Usuário do Portal

Cenários Associados: 6

Papéis: UC

F.1. Informar Usuário

SIGNOS: usuário.nome?

PREVENÇÃO ATIVA: somente usuários cadastrados podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

F.2. Exibir Dados do Usuário

SIGNOS: conjunto(usuário.\*, 0..n)!

Meta G: Alterar Usuário do Portal

Cenários Associados: 6

Papéis: UC

G.1. Informar Usuário

SIGNOS: usuário.nome?

PREVENÇÃO ATIVA: somente usuários cadastrados podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

G.2. Exibir Dados do Usuário

SIGNOS: conjunto(usuário.\*, 0..n)!

G.3. Fornecer Novos Dados do Usuário

SIGNOS: usuário.e-mail?, usuário.endereço?, usuário.cep?, usuário.telefone, usuário.fax, usuário.cnpj/cpf?

TRATAMENTO APOIADO: falta de informações obrigatórias

Meta H: Adicionar Usuário do Portal

Cenários Associados: 2, 3, 5

Papéis: UC

H.1. Fornecer dados do usuário

SIGNOS: usuário.nome?, usuário.e-mail?, usuário.endereço?, usuário.cep?, usuário.telefone, usuário.fax, usuário.cnpj/cpf?

TRATAMENTO APOIADO: falta de informações obrigatórias

Meta I: Consultar Proprietário

Cenários Associados: 10

Papéis: UC, UP

I.1. Informar Proprietário

SIGNOS: proprietário.nome?

PREVENÇÃO ATIVA: somente usuários cadastrados podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

I.2. Exibir Dados do Proprietário

SIGNOS: conjunto(proprietário.\*, 0..n)!

Meta J: Alterar Proprietário

Cenários Associados: 6

Papéis: UC, UP

J.1. Informar Proprietário

SIGNOS: proprietário.nome?

PREVENÇÃO ATIVA: somente usuários cadastrados podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

J.2. Exibir Dados do Proprietário

SIGNOS: conjunto(proprietário.\*, 0..n)!

J.3. Fornecer Novos Dados do Proprietário

SIGNOS: proprietário.endereço?, proprietário.cep?, proprietário.cnpj/cpf

TRATAMENTO APOIADO: falta de informações obrigatórias

Meta K: Adicionar Proprietário

Cenários Associados: 4

Papéis: UC, UP

K.1. Fornecer Dados do Proprietário

SIGNOS: proprietário.nome?, proprietário.endereço?, proprietário.cep?, proprietário.cnpj/cpf?

TRATAMENTO APOIADO: falta de informações obrigatórias

Meta L: Consultar Veículo

Cenários Associados: 9

Papéis: UC, UP

L.1. Informar Veículo

SIGNOS: veículo.RENAVAM?

PREVENÇÃO ATIVA: somente usuários cadastrados podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

L.2. Exibir Dados do Veículo

SIGNOS: conjunto(veículo.\*, 0..n)!

Meta M: Alterar Veículo

Cenários Associados: 9

Papéis: UC, UP

M.1. Informar Veículo

SIGNOS: veículo.RENAVAM?

PREVENÇÃO ATIVA: somente usuários cadastrados podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

M.2. Exibir Dados do Veículo

SIGNOS: conjunto(veículo.\*, 0..n)!

M.3. Fornecer Novos Dados do Veículo

SIGNOS: veículo.placa?, veículo.ano\_fabricação?, veículo.cor?, veículo.tipo?, veículo.chassis?, veículo.marca?, veículo.modelo?, veículo.combustível?, veículo.cnpj/cpf?

TRATAMENTO APOIADO: falta de informações obrigatórias

M.4. Emitir Comprovante

M.4.a.1. Por e-mail

SIGNOS: usuario.e-mail?

M.4.a.2. Por fax

SIGNOS: usuario.fax?

M.4.a.3. Por correio

SIGNOS: usuario.endereço?, usuario.cep?

Meta N: Adicionar Veículo

Cenários Associados: 4

Papéis: UC, UP

N.1. Fornecer Dados do Veículo

SIGNOS: veículo.RENAVAM?, veículo.placa?, veículo.ano\_fabricação?, veículo.cor?, veículo.tipo?, veículo.chassis?, veículo.marca?, veículo.modelo?, veículo.combustível?, veículo.cnpj/cpf?

TRATAMENTO APOIADO: falta de informações obrigatórias

N.2. Emitir Comprovante

N.2.a.1. Por e-mail

SIGNOS: usuario.e-mail?

N.2.a.2. Por fax

SIGNOS: usuario.fax?

### N.2.a.3. Por correio

SIGNOS: usuario.endereço?, usuario.cep?

Meta O: Esclarecer Dúvidas

Cenários Associados: 1, 5

Papéis: U, UC, UP

O.1. Fornecer Critério de Pesquisa

SIGNOS: dúvida.pergunta?

TRATAMENTO APOIADO: falta de informações obrigatórias

O.2. Exibir Resultado

SIGNOS: dúvida.resposta, 0..n!

Meta P: Cadastrar Dúvida

Cenários Associados: 5

Papéis: UC, UP

P.1. Informar Dúvida

SIGNOS: dúvida.pergunta?

PREVENÇÃO ATIVA: somente usuários cadastrados podem executar esse tipo de tarefa

TRATAMENTO APOIADO: falta de informações obrigatórias

P.2. Emitir Comprovante

P.2.a.1. Por e-mail

SIGNOS: usuario.e-mail?

P.2.a.2. Por fax

SIGNOS: usuario.fax?

P.2.a.3. Por correio

SIGNOS: usuario.endereço?, usuario.cep?

Tarefa T1: Efetuar Login

Cenários Associados: 2, 3, 4, 5, 6, 9, 10, 11

Papéis: UC, UP

T1.1. Informar Login

SIGNOS: usuário.nome?

TRATAMENTO APOIADO: login incorreto

## 5.1.3 Modelo de Interação para o IPVA

O modelo de interação complementa as etapas anteriores da MoLIC, pois se preocupa com o processo de comunicação entre as metas no momento da interação. No caso do sistema do IPVA o processo interativo escolhido para representar esta etapa da modelagem é a simulação de uso do pagamento do IPVA para o ano corrente. Como para executar essa tarefa o usuário necessita estar autenticado no sistema, também foi criado o modelo de interação para efetuar o login no portal de voz governamental.

### 5.1.3.1 Modelo Diagramático de Interação

O objetivo do modelo diagramático de interação é o de fornecer uma visão geral do diálogo entre o usuário e o aplicativo como um todo.

#### Cena: Efetuar login

pré: login = Falso

1. u: {fornecer dados de login}
  - p: erro - login inexistente
  - p: repetir ou adicionar login?
  - 1.1. u: {repetir}
    - pós: login = Falso (retornar a cena efetuar login)
  - 1.2 u: {adicionar login}
    - pós: login = Falso (cena adicionar usuário)
2. u: {fornecer dados de login}
  - p: sucesso
  - pós: login = Verdadeiro (retornar a cena principal)
3. u: {abandonar login}
  - pós: login = Falso (cena qualquer)

#### Cena: Pagar o IPVA

pré: login = Verdadeiro

1. u: {optar pelo pagamento do ano atual ou dos anos anteriores}
  - p: erro - opção inexistente
  - pós: pagamento = Falso (retornar a cena pagar IPVA)}
2. u: {optar pelo pagamento do ano atual ou dos anos anteriores}
  - p: sucesso
  - 2.1. u: {ano corrente}
    - 2.1.1. u: {fornecer identificador único do veículo}
      - p: erro - veículo inexistente
      - p: repetir ou adicionar veículo?
    - I. u: {repetir}
      - pós: pagamento = Falso (retornar ao item 2.1.1)
    - II. u: {adicionar veículo}
      - pós: pagamento = Falso (cena adicionar veículo)}
    - 2.1.2. u: {fornecer identificador único do veículo}
      - p: sucesso
      - 2.1.2.1. u: {optar pelo tipo de pagamento}
        - p: erro - opção inexistente
        - pós: pagamento = Falso (retornar ao item 2.1.2.1)
      - 2.1.2.2. u: {optar pelo tipo de pagamento}
        - p: sucesso
    - I. u: {pagamento único com desconto}
      - p: valor total do IPVA com desconto
    - II. u: {pagamento único sem desconto}
      - p: valor total do IPVA sem desconto
    - III. u: {pagamento parcelado}
      - p: valor das três parcelas
  - 2.2. u: {anos anteriores}



- 2.2.1. u: {fornecer identificador único do veículo}
  - p: erro - veículo inexistente
  - pós: pagamento = Falso (retornar ao item 2.2.1)
- 2.2.2. u: {fornecer identificador único do veículo}
  - p: sucesso
- 2.2.2.1. u: {optar pelo tipo de pagamento}
  - p: erro - opção inexistente
  - pós: pagamento = Falso (retornar ao item 2.2.2.1)
- 2.2.2.2. u: {optar pelo tipo de pagamento}
  - p: sucesso
- I. u: {pagamento único sem desconto}
  - p: valor total do IPVA sem desconto
- II. u: {pagamento parcelado}
  - p: valor das três parcelas
- 3. u: {optar pela forma de pagamento}
  - p: erro - opção inexistente
  - pós: pagamento = Falso (retornar ao item 3)
- 4. u: {optar pela forma de pagamento}
  - p: sucesso
- 4.1. u: {imediatos}
  - 4.1.1. u: {fornecer dados da conta bancária}
    - p: erro - conta bancária inexistente
    - pós: pagamento = Falso (retornar ao item 4.1.1)
  - 4.1.2. u: {fornecer dados da conta bancária}
    - p: sucesso
- 4.2. u: {posterior}
  - p: sucesso
- 5. u: {optar pela forma de envio do comprovante}
  - p: erro - opção inexistente
  - pós: pagamento = Falso (retornar ao item 5)
- 6. u: {optar pela forma de envio do comprovante}
  - p: sucesso
  - pós: pagamento = Verdadeiro (cena qualquer)
- 7. u: {abandonar pagamento do IPVA}
  - pós: pagamento = Falso (cena qualquer)

### 5.1.3.2 Estrutura Textual

A estrutura textual, dentro do modelo de interação tem o objetivo de descrever os detalhes do processo de comunicação entre o usuário e o aplicativo.

Cena: Efetuar login

1. Fornecer dados do login

nome? <nome completo do usuário: obrigatório>

Cena: Pagar o IPVA

1. Optar pelo pagamento do ano atual ou dos anos anteriores

opção? <texto 'ano corrente' ou 'anos anteriores': obrigatório>

2. Fornecer identificador único do veículo  
 RENAVAM? <código RENAVAM: obrigatório>

2.1. Ano corrente - Optar pelo tipo de pagamento  
 opção? <texto 'único com desconto', 'único sem desconto' ou 'parcelado': obrigatório>

I. Pagamento único com desconto  
 valor! <valor total do IPVA com desconto>

II. Pagamento único sem desconto  
 valor! <valor total do IPVA sem desconto>

III. Pagamento parcelado  
 data 1! <data da primeira parcela>  
 valor 1! <valor da primeira parcela>  
 data 2! <data da segunda parcela>  
 valor 2! <valor da segunda parcela>  
 data 3! <data da terceira parcela>  
 valor 3! <valor da terceira parcela>

2.2. Anos anteriores - Optar pelo tipo de pagamento  
 opção? <texto 'único sem desconto' ou 'parcelado': obrigatório>

I. Pagamento único sem desconto  
 valor! <valor total do IPVA sem desconto>

II. Pagamento parcelado  
 data 1! <data da primeira parcela>  
 valor 1! <valor da primeira parcela>  
 data 2! <data da segunda parcela>  
 valor 2! <valor da segunda parcela>  
 data 3! <data da terceira parcela>  
 valor 3! <valor da terceira parcela>

3. Optar pela forma de pagamento  
 opção? <texto 'imediate' ou 'posterior': obrigatório>

I. Imediato  
 banco? <código do banco: obrigatório>  
 agência? <número da agência: obrigatório>  
 conta? <número da conta: obrigatório>  
 valor! <valor do débito>

II. Posterior  
 valor! <valor do débito>

4. Optar pela forma de envio do comprovante  
 opção? >texto 'e-mail', 'fax' ou 'correio': obrigatório>

I. E-mail  
 e-mail! <e-mail do usuário>  
 data! <data(s) do pagamento do IPVA>  
 valor! <valor(es) do IPVA a ser(em) pago(s)>  
 proprietário.\*! <dados do proprietário>  
 veículo.\*! <dados do veículo>

II. FAX  
 fax! <número do fax do usuário>  
 data! <data(s) do pagamento do IPVA>  
 valor! <valor(es) do IPVA a ser(em) pago(s)>  
 proprietário.\*! <dados do proprietário>  
 veículo.\*! <dados do veículo>

### III. Correio

endereço! <endereço do usuário>

cep! <número do cep do usuário>

data! <data(s) do pagamento do IPVA>

valor! <valor(es) do IPVA a ser(em) pago(s)>

proprietário.\*! <dados do proprietário>

veículo.\*! <dados do veículo>

## 5.2 Usabilidade no Portal de Voz Governamental

A usabilidade é uma meta fundamental a ser obtida em qualquer processo interativo entre o usuário e um sistema computacional. Entende-se por sistema computacional, todo o tipo de equipamento ou aplicativo que tenha sido desenvolvido com o objetivo de ser manipulado pelo ser humano. A usabilidade para cada um desses sistemas computacionais tem as suas características individuais e dependem principalmente da sua concepção original. No caso do estudo em questão, serão apenas abordados os processos mais relevantes para as interfaces audíveis.

Existe uma grande diferença quanto ao uso, nos aplicativos construídos com foco nos aspectos audíveis, ao serem comparados com outro tipo de aplicativo. Essas particularidades, são significativas tanto na definição dos atributos de usabilidade, como também, nos de acessibilidade. Cabe ressaltar, que nessa análise não serão consideradas as diversas soluções tecnológicas existentes para o desenvolvimento de portais de voz.

As heurísticas para os portais de voz são um grande aliado para alcançar esse objetivo. Sendo assim, serão utilizados os seis grupos de heurísticas (balanceamento da árvore de opções, fraseologia, linguagem, flexibilidade, suporte ao usuário e confiabilidade) já descritas na seção 2, para nortear o seu desenvolvimento.

A construção de aplicativos direcionados para portais de voz devem respeitar todos os fatores de usabilidade, como qualquer outro tipo de aplicativo. De acordo com a norma ISO9241, a usabilidade pode ser expressa por meio da definição dos conceitos de eficácia, eficiência e satisfação.

### 5.2.1 Eficácia

A norma ISO9241 define a eficácia como sendo a capacidade de se atingir a acurácia e a completude. A acurácia é expressa pela capacidade que o usuário tem de resolver o seu problema inteiramente e a completude é expressa pela abrangência que o serviço permite ao usuário resolver os seus problemas.

A eficácia no portal de voz governamental pode ser definida, como a capacidade que o usuário tem de obter o serviço desejado e concluir essa operação com sucesso. Para tal, deve-se utilizar na sua construção, todas as heurísticas existentes, desde que sejam adequadas, para as interfaces audíveis.

As heurísticas utilizadas para tentar tornar o portal de voz eficaz, foram as seguintes:

- a) Balanceamento da árvore de opções: Equilíbrio na quantidade das opções. Observada em vários cenários, entre eles o de 'Pagamento do IPVA';
- b) Fraseologia: Frases objetivas. Observada em vários cenários, entre eles o de 'Informações sobre o IPVA';
- c) Linguagem: Linguagem natural. observada em vários cenários, entre eles o de 'Informações sobre o IPVA';
- d) Flexibilidade: Ausência de entrada de informações (noinput). Observada no cenário 'Informações sobre o IPVA';
- e) Suporte ao Usuário: saída ou ajuda. Observada em vários cenários, entre eles o de 'Solicitação de Ajuda'.

### 5.2.2 Eficiência

A norma ISO9241, define a eficiência como sendo a capacidade de se atingir um determinado objetivo no menor tempo possível. Isso é expresso na otimização de todos os recursos disponíveis (interfaces) para se atingir essa meta.

A eficiência no portal de voz governamental, pode ser definida, como a capacidade que o usuário tem de obter o serviço desejado e concluir essa operação no menor intervalo de tempo possível. Cabe ressaltar, que em um portal de voz, esse tempo deve ser extremamente otimizado, pois a demora na execução das tarefas significa que recursos telefônicos e de

processamento estarão momentaneamente indisponíveis. Como na eficácia, deve-se utilizar na sua construção, todas as heurísticas existentes, desde que sejam adequadas, para as interfaces audíveis.

As heurísticas utilizadas para tentar tornar o portal de voz eficiente, foram as seguintes:

- a) Balanceamento da árvore de opções: Equilíbrio na quantidade das opções. Observada em vários cenários, entre eles o de 'Pagamento do IPVA';
- b) Fraseologia: Frases objetivas. Observada em vários cenários, entre eles o de 'Informações sobre o IPVA';
- c) Linguagem: Linguagem natural. observada em vários cenários, entre eles o de 'Informações sobre o IPVA';
- d) Flexibilidade: Acesso rápido as opções (atalho). Observada em vários cenários, entre eles o de 'Atalho para os Serviços do IPVA'
- e) Suporte ao Usuário: Corrigir falhas. Observada em vários cenários, entre eles o de 'Cadastramento de Usuário'.

### 5.2.3 Satisfação

A norma ISO9241, define a satisfação como sendo o conjunto de sensações positivas e negativas que um usuário tem ao usar um determinado aplicativo. Esse sentimento pode ser expresso pelo grau de rejeição, ou não, que o usuário tem ao utilizá-lo.

A satisfação no portal de voz governamental, pode ser definida, como a capacidade que o usuário tem de executar as suas tarefas com facilidade e rapidez; bem como, na necessidade de obter serviços similares, desejar reutilizá-lo, ou mesmo, recomendá-lo para outros usuários. Outro aspecto relevante é o processo de autenticação no portal, que é executado uma única vez e implementa a técnica de análise biométrica da voz humana. Como nos itens anteriores, deve-se utilizar na sua construção, todas as heurísticas existentes, desde que sejam adequadas, para as interfaces audíveis.

As heurísticas utilizadas para tentar tornar o portal de voz um aplicativo que provoque satisfação ao usuário, foram as seguintes:

- a) Balanceamento da árvore de opções: Equilíbrio na quantidade das opções. Observada em vários cenários, entre eles o de 'Pagamento do IPVA';
- b) Fraseologia: Frases objetivas. Observada em vários cenários, entre eles o de 'Informações sobre o IPVA';
- c) Linguagem: Linguagem natural. observada em vários cenários, entre eles o de 'Informações sobre o IPVA';
- d) Flexibilidade: Acesso rápido as opções (atalho). Observada em vários cenários, entre eles o de 'Atalho e Cadastramento de Dúvidas';
- e) Suporte ao Usuário: saída ou ajuda. Observada em vários cenários, entre eles o de 'Solicitação de Ajuda'.
- f) Confiabilidade: Integridade das informações. Observada em vários cenários, entre eles 'IPVA para veículo novo'.

Em suma, a eficiência, a eficácia e a satisfação do usuário são metas a serem atingidas no portal de voz. O fato de aplicar os conceitos acima descritos, não garante plenamente o seu sucesso. Portanto, mais uma vez, a recomendação é que sejam realizados testes de usabilidade para avaliar essas implementações, proporcionando um processo de melhora na qualidade do aplicativo de forma consistente.

### 5.3 Contribuição da MoLIC no Portal de Voz Governamental

A MoLIC, por implementar técnicas da engenharia semiótica e conceitos de IHC, auxiliou de forma fundamental a definição de alguns atributos de usabilidade para o portal de voz governamental. É notório, que ela não foi concebida para modelar especificamente aplicações audíveis, porém, a sua adaptação para esse tipo de funcionalidade foi algo relativamente simples. Os seus elementos como os cenários, o modelo de tarefas e o modelo de interação foram muito úteis durante o processo de modelagem do portal de voz.

Os cenários se transformaram em um grande aliado na definição de alguns atributos de usabilidade. A necessidade de descrever todos os diálogos e interações, permitiu que algumas falhas nesses atributos fossem identificadas e corrigidas. Situações de erro, repetição de textos interativos, redundância nos diálogos, apoio ao usuário em determinados momentos, entre outros, foram detectados com antecedência. Sendo assim, heurísticas como o balanceamento

da árvore de opções, fraseologia, suporte ao usuário e flexibilidade, puderam ser melhoradas nessa etapa.

O modelo de tarefas, na etapa de definição das metas do usuário, permitiu identificar o principal foco no uso de determinadas tarefas. Dessa forma, os processos se tornaram mais objetivos e otimizados. Já na etapa do modelo diagramático de tarefas, a integração dos elementos que compõem a MoLIC, como signos, papéis e cenários, fornecem uma visão mais global das tarefas, favorecendo o diagnóstico de possíveis problemas. Heurísticas como suporte ao usuário e flexibilidade, também puderam ser melhoradas nessa etapa.

Outra contribuição do modelo de tarefas foi a possibilidade de identificar o reuso de algumas tarefas ou estereótipos, como o logon do usuário no sistema, que permitiu a construção de um modelo mais otimizado. Também, a utilização de tarefas ubíquas no modelo, como sair do portal ou falar com um atendente, em qualquer ponto do processo interativo, proporcionou maior flexibilidade ao usuário. Por fim, a necessidade do apoio e tratamento de erros, observada em várias tarefas, forneceu subsídios para a criação de um modelo mais consistente.

O modelo de interação, teve apenas como alvo as cenas de autenticação no portal e pagamento do ano corrente do IPVA. A complexidade do modelo diagramático de interação e da estrutura textual, não permitiu que o sistema do IPVA fosse contemplado como um todo. Entretanto, essas duas cenas são suficientes para demonstrar a importância dessa etapa no processo de modelagem. Esses dois elementos foram muito úteis para identificar redundâncias nos processos comunicativos e nos diálogos, tornando possível minimizar a quantidade de interações. Heurísticas como flexibilidade, fraseologia e confiabilidade, também puderam ser melhoradas nessa etapa.

A MoLIC não contempla diretamente dois itens de heurística considerados fundamentais na construção de um portal de voz. Porém, na criação dos cenários e do modelo de tarefas, durante a definição dos diálogos, os itens de heurística denominados fraseologia e linguagem natural, passaram a ser considerados no processo de comunicação entre o usuário e o aplicativo. Isso significa que a MoLIC não aborda formalmente esses itens, mas fornece a oportunidade do projetista analisar esses dois elementos com alguma antecedência.

Dessa maneira, os signos necessitaram de uma revisão de conceitos, ao longo da criação de todos os artefatos propostos pela linguagem. O que se pôde observar é que nas interfaces audíveis os objetos de diálogo, frases utilizadas no processo de comunicação entre o aplicativo e o usuário, isto é, fraseologia, também devem ser considerados como um signo. Esses signos, que devem implementar a linguagem natural, como sendo a principal forma de comunicação, foram obtidos em uma versão final da fraseologia para o portal de voz governamental, que pode ser exemplificada, no cenário a seguir:

### Cenário 3 - Pagamento do IPVA

*PVG: "Qual é o serviço desejado?"*

Mariana: "Pagar o IPVA"

*PVG: "Favor fale o nome completo" {11}*

Mariana: "Mariana Alves"

*PVG: "Usuário não encontrado no cadastro"*

*"Caso já tenha executado o cadastro, basta dizer o seu nome completo novamente. Caso contrário, diga Adicionar usuário" {9,13}*

Mariana: "Adicionar usuário"

*PVG: "Favor fale o nome completo"*

Mariana: "Mariana Alves"

*PVG: "Favor fale o endereço completo"*

Mariana: "Rua das Pombas, 136"

*PVG: "Digite através do teclado o número do CEP"*

*"Ddignite através do teclado o número do CPF ou do CNPJ"*

*"Favor fale o e-mail"*

Mariana: "malves@ig.com.br"

*PVG: "Digite através do teclado o número do FAX"*

*PVG: "Os seus dados cadastrais são:"*

*nome completo é Mariana Alves"*

*"O endereço completo é Rua das Pombas, 136"*

*"O número do CEP é 03096 010"*

*"O número do CPF ou do CNPJ é 23.456-7"*

*"O e-mail é malves@ig.com.br"*

*"O número do FAX é 2345-6789"*

*"Deseja ouvir novamente os seus dados cadastrais?"*

Mariana: "não"

*"Os dados cadastrais estão corretos?"*

Mariana: "sim"

*PVG: "Cadastramento executado com sucesso"*

*PVG: "Deseja fazer o pagamento deste ano ou dos anos anteriores?" {23,24}*

Mariana: "deste ano"

*PVG: "Deseja fazer pagamento único ou parcelado?" {18}*

Mariana: "pagamento único"

*PVG: "O desconto somente será aplicado para os pagamentos realizados até o dia 08 de Janeiro de 2006" {30}*

*"Digite através do teclado o código RENAVAL"*

*PVG: "O valor a ser pago é de R\$ 500,00"*

*"Deseja a emissão de um boleto bancário ou pagar agora?" {19,20}*



Mariana: "Hummm! ... pagar agora"

PVG: "*Informe o número do banco e da agência*"

Mariana: "151 09369"

PVG: "*Desculpe, não entendi. Favor repetir*"

Mariana: "Banco 151 e agência 09369"

PVG: "*Favor digitar através do teclado o número da conta*"

PVG: "*Favor digitar através do teclado a senha*"

PVG: "*Confirme, o valor a ser debitado da sua conta é de R\$ 500,00?*"

Mariana: "sim"

PVG: "*Deseja receber o comprovante através de e-mail, fax ou correio?*" {12,25}

Mariana: "correio"

PVG: "*Operação concluída, deseja sair ou acessar outro serviço?*"

Cabe ressaltar que a fraseologia proposta, somente atinge uma maior consistência e maturidade, após a conclusão das três etapas recomendadas pela MoLIC. Contudo, isso não invalida a necessidade desse conjunto de diálogos passar por um teste de usabilidade formal, conforme é recomendado aos projetos que desejam ser desenvolvidos dentro dos conceitos de IHC.

Após a conclusão das etapas propostas pela MoLIC, o conjunto de artefatos obtidos serão utilizados nas próximas fases do projeto, auxiliando de forma significativa no desenvolvimento e na implantação de sistemas computacionais. Por outro lado, existem alguns aspectos negativos na sua implementação. As etapas de modelagem, cenários, modelo de tarefas e modelo de interação, são muito trabalhosas e demandam um grande esforço na sua concepção. Portanto, o tempo e o custo de implantação da MoLIC deve ser um ponto relevante a ser considerado no planejamento do projeto.

Esta seção apresentou a modelagem do IPVA para o portal de voz governamental, através da MoLIC e também abordou alguns aspectos da sua implementação. A próxima seção fará algumas considerações sobre as contribuições desse estudo e analisará as perspectivas para trabalhos similares no futuro.

## 6. CONCLUSÃO

As interfaces audíveis provaram ser grandes aliadas no desenvolvimento de aplicativos que favorecem a acessibilidade. O estudo em questão, ao longo das suas sessões, forneceu diversos subsídios que comprovam essa afirmação.

A seção 2 abordou vários fatores em que o som, pode auxiliar no desenvolvimento de excelentes interfaces computacionais. Aspectos como a autenticação biométrica através da voz humana, o crescimento tecnológico observado nos processos de reconhecimento e reprodução da fala, as heurísticas identificadas para os portais de voz e os estudos sobre as interfaces audíveis de fala e de não fala, forneceram atributos bastante úteis no desenvolvimento dos aplicativos.

A seção 3 forneceu uma visão geral dos processos de modelagem para sistemas interativos e detalhou os principais conceitos daMoLIC, que permitiram criar uma base de conhecimento satisfatória para que no momento de modelar o IPVA no portal de voz governamental, fossem utilizados todos os recursos disponíveis para construir uma interface acessível e de fácil uso.

A seção 4 permitiu identificar as dificuldades encontradas durante o processo de análise e levantamento de requisitos, necessários para migrar o atual portal de governo eletrônico para um portal de voz. Os vários problemas detectados, foram em sua maioria, provocados pelas diferenças tecnológicas das interfaces visuais, ao serem comparadas com as audíveis. Porém, a MoLIC serviu como um grande agente facilitador para identificar e sanar os problemas encontrados.

A seção 5 implementou a MoLIC, para o IPVA dentro do portal de voz governamental. Esse processo, propiciou implementar de forma integral as três etapas propostas na MoLIC. Os cenários, o modelo de tarefas do usuário e o modelo de interação forneceram recursos para dirimir alguns problemas de interação. Já as heurísticas para os portais de voz, em conjunto com a MoLIC, permitiu trabalhar no sentido de corrigir problemas de comunicação, identificadas nos diálogos entre o usuário e o aplicativo.

Baseado nessas afirmações, pode-se concluir que as interfaces audíveis, surgem como uma excelente alternativa na construção de aplicativos que privilegiam a usabilidade. A

contribuição das Interfaces Audíveis na inclusão digital, passam a ser um fator importante para a tomada de decisão por parte dos projetistas. As interfaces visuais, apesar de todos os esforços do poder público com a criação do Modelo de Acessibilidade de Governo Eletrônico (e-MAG), a lei de acessibilidade citado em Brasil (2004), e iniciativas por entidades como a WORLD Wide Web Consortium (2003) com o *Web Content Accessibility Guidelines 2.0*, tem provado ser muito difíceis de se tornarem uma realidade.

As recomendações propostas nem sempre são de fácil implementação. Os procedimentos a serem adotados, exigem um conjunto de técnicas específicas e um treinamento especial para que os desenvolvedores desses sítios, possam aplicar esses conceitos de acessibilidade. Outro ponto que deve ser levado em consideração, é a dificuldade que certas camadas da população tem para acessar a Internet. Sejam elas, por motivos sócio-culturais, ou provocados por fatores limitantes, como usuários portadores de necessidades especiais, analfabetos, idosos e outros.

As interfaces audíveis, apesar de não serem o principal foco dessas iniciativas, podem ser um excelente agente facilitador da inclusão digital, pois utilizam a fala como a principal forma de interação entre o usuário e o computador, que é o processo de comunicação mais natural do ser humano. Situações de auto-atendimento, onde a tarefa é executada diretamente pelo usuário, sem a intervenção de um atendente, poderiam ser melhoradas de forma significativa, com a utilização desse tipo de interface. De tal forma, que proporcionariam maior agilidade nas atividades a serem realizadas.

Serviços oferecidos pelo Governo do Estado de São Paulo, como o Poupatempo, que atende uma enorme quantidade de cidadãos, das mais diversas classes sociais, com cultura, experiência e idade muito heterogêneas, são um bom exemplo de como esse tipo de interface quando utilizadas, podem ser úteis. O uso da voz humana como forma de interação seria extremamente natural. Os processos de comunicação se tornariam mais ágeis, principalmente para as pessoas portadoras de necessidades especiais, que hoje não tem nenhuma opção concreta de uso; bem como pessoas idosas com problemas de visão reduzida e as pessoas analfabetas, que ao utilizarem uma interface de comunicação direta, se tornariam mais independentes. Por sua vez, essa independência na utilização desses serviços, trarão um grande benefício no que se refere a auto-estima do cidadão e uma provável diminuição na demora no atendimento nesses locais.

## 6.1 Dificuldades Identificadas

Conforme abordado anteriormente, os procedimentos de conversão de um aplicativo concebido com foco nos aspectos visuais para um audível, passam por um grande conjunto de processos de adequação e de adaptação. No caso do portal de governo eletrônico, as coisas não foram diferentes.

Inicialmente, houve a necessidade de identificar as diferenças fundamentais nos processos comunicativos das duas interfaces e propor soluções alternativas para os problemas encontrados. Durante esse processo, surgiram várias perguntas e as suas respostas tiveram que ser obtidas baseadas nas soluções tecnológicas disponíveis. Alguns desses questionamentos foram: como fazer para substituir a imagem pelo som? Como fornecer as informações, ou as opções, sem sobrecarregar o usuário? Como utilizar a linguagem para facilitar o processo de comunicação? Como gerar um documento formal da operação realizada? Como permitir um acesso rápido ao serviço desejado? Entre outras.

No caso da pergunta de "Como fazer para substituir a imagem pelo som?" a resposta foi óbvia. Ambos os processos comunicativos tem as suas próprias características que os individualizam. A imagem tem o poder de transmitir uma grande quantidade de informação de uma só vez e o seu aspecto visual facilita a memorização e a navegação ao longo do processo comunicativo, entre o usuário e o aplicativo. Já o som é algo mais abstrato, facilita a dispersão e prejudica a memorização. Esses aspectos negativos foram analisados e a solução adotada, foi a de tornar os diálogos interativos o mais natural possível, proporcionando ao usuário a sensação de interagir, ao longo do processo comunicativo, diretamente com um atendente. O aplicativo também permite ao usuário retornar, sair e falar com um atendente a qualquer momento. Essas opções tem o objetivo de facilitar a navegação e suprir a dificuldade de memorização.

A pergunta "Como fornecer as informações, ou as opções, sem sobrecarregar o usuário?" também foi analisada e a sua solução veio através do uso das heurísticas para portais de voz. A interface visual, em uma única tela, pode fornecer uma grande quantidade de opções de serviço, sem sobrecarregar o usuário. As interfaces audíveis não podem utilizar esse mesmo método que sobrecarregariam os usuários, provocando a dispersão e o cansaço. Para

minimizar esses impactos, a solução adotada foi a de construir árvores de opções mais objetivas e claras. Este estudo não é elementar, pois houve a necessidade de analisar todo o fluxo de opções atualmente existentes e adaptá-las para um novo fluxo, com a preocupação de implementar o balanceamento da árvore de opções

Já a pergunta de "Como utilizar a linguagem para facilitar o processo de comunicação?" foi resolvida com o uso da linguagem natural, que é outra heurística. A imagem por si, fornece muita informação e os textos, quando utilizados, tem o único objetivo de auxiliar na comunicação. O som, ao contrário, está totalmente baseado na fala para realizar esse processo comunicativo. Portanto, o uso da linguagem natural, sem abreviações, termos técnicos e palavras complexas, propiciaram a criação de diálogos interativos muito eficientes.

Outra pergunta que surgiu foi "Como gerar um documento formal da operação realizada?", o aplicativo atual permite através de uma impressora associada ao computador, imprimir um documento com o resultado da operação executada. No caso do novo aplicativo, isso não seria possível. Sendo assim, no novo aplicativo foi fornecida a possibilidade de enviar esse documento por e-mail, fax ou correio. Para tal, houve a necessidade de solicitar esses dados no momento do cadastramento do usuário do portal de voz.

Por fim, "Como permitir um acesso rápido ao serviço desejado?" essa característica tem como objetivo principal, diminuir a possibilidade do cansaço e da dispersão dos usuários. A tecnologia de reconhecimento da voz humana, permite ditar frases que abrangem várias opções ao mesmo tempo. Esta funcionalidade, é particularmente muito útil, para usuários experientes. Essa solução também, foi adotada para tentar suprir algumas desvantagens de acesso observadas no portal de voz, quando comparadas com as do portal visual. Resumindo, houve a necessidade de fazer uma análise bastante detalhada do portal de governo eletrônico do estado de São Paulo, para adaptá-lo ao portal de voz proposto.

## 6.2 Contribuição da MoLIC

A utilização de uma modelagem interativa, em especial a MoLIC, foi extremamente importante para identificar que haviam falhas nos diálogos inicialmente propostos e forneceu instrumentos para tentar corrigi-los. No momento da construção dos cenários, houve a

possibilidade de obter uma visão geral dos processos interativos que seriam executados pelo usuário, ao utilizar o portal de voz para acessar o IPVA. Nesta etapa foi possível verificar, a grosso modo, que a migração dos atuais serviços existentes no portal tradicional, seria uma tarefa bastante complexa.

A primeira versão dos cenários criados para o portal de voz, praticamente não trouxe muitas novidades, já que eles tem o objetivo de apresentar um conjunto de atividades a serem executadas pelo usuário, sem definir os objetos utilizados nesse diálogo. A segunda versão que foi concebida com um maior nível de detalhamento, possibilitou construir diálogos mais próximos da linguagem natural, que era o objetivo a ser atingido. Após isso, os requisitos de interação e os objetos do diálogo se confundiram, gerando cenários com uma característica diferente da proposta pela MoLIC, com um foco muito grande nos objetos do diálogo.

Infelizmente, esse é um erro que pode ocorrer com muita frequência ao se modelar interfaces audíveis. Se o projetista não ficar atento na criação desses diálogos, que são caracterizados por uma comunicação direta, através da fala entre o usuário e o aplicativo, ele irá incidir neste erro conceitual. Já na terceira e última versão dos cenários, foram implementadas todas as experiências anteriores e foi gerado um conjunto de situações descritas de uma forma mais consistente. Este fato ajudou significativamente na identificação de algumas falhas no processo de comunicação. Entretanto, ainda havia muito a ser melhorado no projeto.

O segundo artefato da MoLIC, que é o modelo de tarefas do usuário, confirmou essa desconfiança. Quando foram definidas as metas do usuário ao utilizar o IPVA, que são os objetivos a serem atingidos pelo usuário ao utilizar o aplicativo, foram observadas várias lacunas no projeto interativo. A descrição minuciosa dessas metas e de suas respectivas tarefas associadas, permitiram fazer uma análise detalhada de cada diálogo interativo, proporcionando um produto com menos falhas na comunicação e uma significativa redução nas etapas para se atingir uma determinada tarefa. Um caso típico pôde ser observado na tarefa de logon no sistema, onde no portal visual é solicitada em um determinado momento do processo interativo, e já no portal de voz, passou a ser solicitado em um outro momento e apenas uma vez.

Por fim, o último artefato que é o modelo de interação, forneceu através do modelo diagramático de interação e da estrutura textual, subsídios para consolidar os diálogos

interativos entre o usuário e o aplicativo do IPVA. Nesta etapa foi possível identificar as ligações entre as metas obtidas no modelo de tarefas. Dessa maneira foram suprimidas as redundâncias e consolidadas as interações do pagamento do IPVA para o ano corrente.

Cabe ressaltar, que a implementação dos artefatos da MoLIC no projeto do portal de voz governamental, demonstrou ser uma tarefa bastante trabalhosa. É provável que esses procedimentos quando executados por um profissional mais experiente, nesse tipo de ferramenta, possam implementar os elementos da MoLIC com maior facilidade.

Outra coisa que cabe ser observada, é que ao se seguir a risca os conceitos da MoLIC, itens como a fraseologia e a linguagem natural, que são considerados fundamentais para os portais de voz, não foram plenamente contemplados ao longo da modelagem. Esse é um aspecto muito importante a ser abordado, pois o processo de adequação da MoLIC para modelar sistemas audíveis, fica com essa grande lacuna a ser preenchida. Uma forma de atenuar essa limitação é considerar a fraseologia em um portal de voz, como sendo um signo, conforme apresentado na seção 5.

### 6.3 Implementação do Portal de Voz Governamental

A obtenção de todos os artefatos resultantes da MoLIC, forneceu dados suficientes para a construção de um projeto de interação mais consistente. Entretanto, isto não garante o sucesso na sua implementação. Portanto, torna-se necessário criar esses processos comunicativos, baseados nas informações obtidas através da MoLIC e aperfeiçoá-las. A recomendação é a de utilizar protótipos contendo esses diálogos interativos e executar testes de usabilidade com um grupo de usuários que representem de forma significativa o público alvo do portal de voz governamental. Somente após a validação dessa fase do projeto, as próximas etapas do ciclo de vida do projeto deverão ser executadas.

Outro aspecto a ser considerado, no momento da implementação é o da limitação tecnológica. Casos como o observado no cenário 3, onde o usuário fornece o e-mail através da fala, deverão ser melhor avaliados. Experiências têm demonstrado que conteúdos complexos contendo caracteres especiais tem um baixo nível de reconhecimento. Para atenuar esse problema, algumas alternativas podem ser utilizadas, como o de soletrar o texto do conteúdo,

ditando letra por letra da informação desejada, ou fazer uso do legado desse dado, através de um cadastro existente.

Já a seção que aborda as linguagens Voice XML e SALT, faz uma análise comparativa entre ambos os padrões, e recomenda que alguns fatores culturais da organização e de mercado devem ser considerados no momento da opção de qual linguagem será escolhida como padrão de desenvolvimento. No caso de aplicações governamentais o Voice XML leva uma ligeira vantagem, pois atualmente o padrão para o desenvolvimento de soluções de TI (Tecnologia da Informação) no poder público, é o de utilizar plataformas de desenvolvimento gratuitas, como o Java. Contudo, existem situações, mesmo no poder público, que o SALT se enquadraria perfeitamente, dada a cultura de desenvolvimento baseada em .NET da Microsoft.

O portal de voz governamental do Estado de São Paulo, em especial o da Secretaria dos negócios da fazenda, poderá ser desenvolvido em ambos padrões de linguagem. Existe cultura de desenvolvimento tanto para o Java (Voice XML) como para o SALT (.NET). Porém, particularmente, considero muito interessante a implantação em uma solução totalmente baseada no Java, com o uso do Voice XML.

A plataforma de desenvolvimento do portal de voz governamental, da mesma forma que os padrões de linguagem, poderão ser escolhidos entre o da IBM com o WebSphere, o da Oracle com o AS10G ou o da Microsoft com o .NET. Neste caso, a opção tecnológica mais adequada é a do WebSphere, dada a característica das aplicações existentes da Secretaria da Fazenda.

Cabe ressaltar, que o objetivo do estudo neste aspecto, foi o de apresentar esse conjunto de ferramentas para fornecer uma visão geral da solução como um todo. Sendo assim, este trabalho não visa entrar nos detalhes do processo de implementação de um portal de voz.

#### 6.4 Trabalhos Futuros

O estudo também oferece a oportunidade para que em trabalhos similares ou futuros, sejam implementados esses mesmos conceitos acima referenciados. No caso do portal de voz com foco em governo eletrônico, seria muito interessante implantar esse modelo proposto de forma efetiva, já que ele atingiu o seu objetivo. Para projetos similares, a experiência aqui relatada,



pode representar um ponto de partida e de incentivo aos projetistas, com o objetivo de implementarem, além das técnicas de modelagem propostas pela MoLIC, as recomendações obtidas para as interfaces audíveis, em aplicativos que utilizem o som como a principal forma de comunicação.

## REFERÊNCIAS

APLICAÇÕES com Reconhecimento de Voz. Disponível em :

<[http://www.malima.com.br/article\\_read.asp?id=175](http://www.malima.com.br/article_read.asp?id=175)>. Acesso em: 25 jul. 2006

BARBOSA, S.D.J.; ; PAULA, M.G. Designing and Evaluating Interaction as Conversation: a Modeling Language based on Semiotic Engineering. In: **Interactive Systems. Design, Specification, and Verification**. [S.l.]: Springer Verlag, 2003. p. 16 – 33. (ISBN: 3-540-20159-9) Disponível: <<http://www.serg.inf.puc-rio.br>>. Acesso em: 13 dez. 2004.

BARBOSA, S.D.J.; DE SOUZA, C.S. ; PAULA, M.G. The Semiotic Engineering Use of Models for Supporting Reflection-In-Action. Proceedings of HCI International 2003. Crete, Greece. Disponível em: <<http://www.serg.inf.puc-rio.br>>. Acesso em: 13 dez. 2004.

BARRASS, S. **Auditory Information Design**.. 349f. Tese (Doutorado), Australian National University. 1998 Disponível em: < <http://thesis.anu.edu.au/public/adt-ANU20010702.150218/index.html> > . Acesso em 14 Jul., 2006.

BLATTNER, M.M.; PAPP A.L.; GLINERT, E.P. Sonic Enhancement of Two-Dimensional Graphics Displays, in KRAMER G. (ed) Auditory Display : Sonification, Audification and Auditory Interfaces, SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII, Addison-Wesley Publishing Company, Reading, MA, U.S.A.. 1994

BLATTNER, M.; SUMIKAWA D.; GREENBERG R. (1989) Earcons and Icons: Their Structure and Common Design Principles, Human Computer Interaction, 4, 1, Lawrence Erlbaum Associates, London.

BLY, S. **Sound and computer information presentation**. 1982. Não paginado. Tese (Doutorado), University of California, Davis, California, USA, 1982. (Unpublished doctoral dissertation, Lawrence Livermore National Laboratory and University of California, Davis, CA - . Thesis No. UCRL53282)

BRANCO, G. ; ALMEIDA L.; BEIRES, N. et al. Evaluation of a multimodal Virtual Personal Assistant. In: International Symposium on Human Factors in Telecommunication, 20, 2006. Sophia-Antipolis, France. **Proceedings...** France:HFT, 2006. Disponível em: < [http://www.hft.org/HFT06/paper06/35\\_Branco.pdf](http://www.hft.org/HFT06/paper06/35_Branco.pdf)>. Ver também: < [http://portal.etsi.org/docbox/Workshop/HFT2006/session7/35\\_Branco/35\\_Branco.ppt](http://portal.etsi.org/docbox/Workshop/HFT2006/session7/35_Branco/35_Branco.ppt) >. Acesso em : 11 ago. 2006

BRASIL. Decreto lei 5296, de 2 de dezembro de 2004. Lei de acessibilidade. Disponível em: < <http://www.acessobrasil.org.br/index.php?itemid=43>>. Acesso em: 25 jul. 2006.

BREWSTER, S.; LUMSDEN, J.; BELL, M et al. **Multimodal ‘Eyes-Free’ Interaction Techniques for Wearable Devices**. Glasgow : University of Glasgow / Faculty of Information and Mathematical Systems , 2003. Disponível em: < <http://www.dcs.gla.ac.uk/~stephen/papers/CHI2003.pdf> . Acesso em: 25 jan. 2004.

BREWSTER, S.A . Providing **structured method for integrating non-speech audio into humancomputer interfaces**.1994. 292 f. Tese (Doutorado), University of York, UK, 1994. Disponível em : <[http://www.dcs.gla.ac.uk/~stephen/papers/theses/Brewster\\_thesis.pdf](http://www.dcs.gla.ac.uk/~stephen/papers/theses/Brewster_thesis.pdf) >Acesso em 14 jul., 2006

BURKHARDT, F.; STEGMANN, J.; VAN BALLEGOOY, M. **A Voiceportal Enhanced by Semantic Processing and Affect Awareness**. (T-Systems International GmbH) Disponível em: < <http://felix.syntheticspeech.de/publications/enhancingVoiceportals.pdf> >.Acesso em: 11/03/2006

CISCO SYSTEMS INC.; et al. **Speech Application Language Tags (1.0 Specification)**. [S.l:s.n], 2002. Disponível em : <http://www.saltforum.org/saltforum/downloads/SALT1.0.pdf> . Acesso em: 26 jul. 2006

CONSTANTINE, Alexandre. **A era da Voz**. Espírito Santo : SUCESU-ES, 2003. Disponível em: [http://www.sucesues.org.br/documentos/index.asp?cod\\_noticia=378](http://www.sucesues.org.br/documentos/index.asp?cod_noticia=378). Acesso em 14 jul., 2005

DEPARTMENT OF DEFENSE - **MIL-STD-1472F**. Design Criteria Standard Human Engineering .EUA : DEPARTMENT OF DEFENSE, 1999. Disponível em: < <http://hftag.dtic.mil/docs-hfs/mil-std-1472f.pdf>>. Acesso em: 19 jul., 2006.

ECO, U. **Tratado Geral de Semiótica**. São Paulo: Perspectiva, 1976.

FONSECA, H.; SANTOS, V.; FERREIRA, A . **A audição: constituição do sistema auditivo humano**, 2002. Disponível em: <[http://telecom.inescn.pt/research/audio/cienciaviva/constituicao\\_audicao.html](http://telecom.inescn.pt/research/audio/cienciaviva/constituicao_audicao.html) >. Acesso em: 30 Jul. 2006.

GAVER, W.W. Using and Creating Auditory Icons. In: Kramer G. (Ed.). Auditory Display : Sonification, Audification and Auditory Interfaces. USA: Addison-Wesley, 2004. p.417-446. (Proceedings Volume 18, Santa Fe Institute Studies in the Sci). Disponível em: <[http://72.14.209.104/search?q=cache:ilIq9XiLLsUJ:www.interaction-design.org/references/authors/william\\_w\\_gaver.html+%22Auditory+Icons%22+%221986%22&hl=pt-BR&gl=br&ct=clnk&cd=10](http://72.14.209.104/search?q=cache:ilIq9XiLLsUJ:www.interaction-design.org/references/authors/william_w_gaver.html+%22Auditory+Icons%22+%221986%22&hl=pt-BR&gl=br&ct=clnk&cd=10)>. Acesso em: 19 Set 2006.

GAVER, W.W. **Auditory Icons**, Human-Computer Interaction, 2, 2, Lawrence Erlbaum Associates, London. 1986 Disponível em:  
<<http://www.billbuxton.com/AudioUI06icons.pdf#search=%22%22Auditory%20Icons%22%22>> Acesso em: 19 Set 2006.

GRUENSTEIN, A . **Automatic Grammar Construction** . Cambridge-EUA : MIT, 2002. Disponível:< <http://www.mit.edu/~alexgru/ssp115.pdf> >.Acesso em : 25 maio 2006

HONE, K.S.; GRAHAM, R. Subjective assessment of speech-system interface usability. In: European Conference on Speech Communication and Technology, 7, 2001. Denmark. **Proceedings...** Aalborg, Denmark : ISCA, 2001. p.2083-2086. [http://www.isca-speech.org/archive/eurospeech\\_2001/e01\\_2083.html](http://www.isca-speech.org/archive/eurospeech_2001/e01_2083.html)

Inovação Tecnológica (2002) Projeto Oxygem revela novas tecnologias para a informática Disponível em:  
<http://www.inovacaotecnologica.com.br/noticias/noticia.php?artigo=010150020724> Acesso em: 25 Ago. 2006.

INTERNATIONAL STANDARD. ISO 9241-11(E). **Ergonomic requirements for office work with visual display terminals (VDTs)-part 11: Guidance on usability**. [s.l.] 1998

KONDRATOVA, I. **Performance and Usability of VoiceXML Application**. Canada : NRC, 2004. Disponível em : < <http://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-47154.pdf>. Acesso em: 27 jul. 2006

LUCAS, P. An Evaluation of the Communicative Ability of Auditory Icons and Earcons. In: International Conference on Auditory Display, 2, 1994. Santa Fé -New Mexico, USA. Proceedings... USA : [S.l.], 1999. Disponível em :  
<http://www.icad.org/websiteV2.0/Conferences/ICAD94/papers/Lucas.pdf> . Acesso em: 12 ago. 2006.

MENCONI, D. A voz do dono e o dono da voz. **Isto è digital**. Edição especial [2004?]. Disponível em: <<http://www.zaz.com.br/istoe/digital/tecnologia.htm>> . Acesso em: 30 Jul. 2006.

MOORE, B.C. **An introduction to the psychology of hearing**. 4 ed. London:Academic Press,1997.

NIELSEN, J. **Usability Engineering**, Academic Press, 1993.

OLIVEIRA, O.L. **Design da Interação em Ambientes Virtuais: uma abordagem semiótica**. 2000. 190f. Tese (Doutorado), Instituto de Computação, Universidade Estadual de Campinas, Campinas-SP., 2000.

PAKUCS, B. **Butler**: Universal Speech Interface for Mobile Environments. Stockholm, Sweden : Royal Institute of Technology, [19--?] . Disponível: [www.speech.kth.se/~botte/publications/mobhci\\_04.pdf](http://www.speech.kth.se/~botte/publications/mobhci_04.pdf) Acesso em: 30 Jul. 2006

PAULA, M. G.: **Projeto da interação humano-computador baseado em modelos fundamentados na engenharia semiótica**: construção de um modelo de interação, Dissertação de mestrado, Departamento de Informática, PUC-Rio, 2003.

PROAKIS, J.; DELLER, J.; HASSEN, J.H.L. **Discrete-time processing of speech signals**. New Jersey: Prentice-Hall,1993.

ROBINSON T. Speech Analysis. 1996. Disponível em < <http://svr-www.eng.cam.ac.uk/~ajr/SpeechAnalysis/> >. Acesso em: 25 jul., 2006

SANTAELLA, L. **O que é Semiótica**. 12.ed.. São Paulo: Editora Brasiliense, 1996

SÃO PAULO(Estado). Governo do Estado de São Paulo. **Portal do Cidadão do governo do estado de São Paulo**. Disponível em < <http://www.cidadao.sp.gov.br> >. Acesso em 26 nov. 2006.

SILVA, B. S.; BARBOSA, S. D. J. **Modelando a interação do NiTA: um estudo de caso e extensões ao MoLIC**. In: SIMPÓSIO SOBRE FATORES HUMANOS EM SISTEMAS COMPUTACIONAIS - MEDIANDO E TRANSFORMANDO O COTIDIANO, 6.,2004., Curitiba. Anais..., Curitiba:UFPR; CEIHC, 2004. p.181-184. Disponível em : <<http://www.serg.inf.puc-rio.br/ihc/papers/IHC2004/181-184-IHC2004-.pdf>>. Acesso em: 26 jul. 2006.

SILVA, B.S. **MoLIC Segunda Edição: revisão de uma linguagem para modelagem da interação humano-computador**. 2005. Dissertação (Mestre em informática) ,Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2005.

VALIATI, J.F. **Reconhecimento de Voz para Comandos de Direcionamento por meio de Redes Neurais**. Disponível em: <<http://www.inf.ufrgs.br/pos/SemanaAcademica/Semana99/joaofrancisco/joaofrancisco.html>>. Acesso em: 09 ago. 2006.

WORLD WIDE WEB CONSORTIUM. *Voice eXtensible Markup Language (2.0 Specification)*. [S.l.]:MIT;ERCIM; Keio, 2004. Disponível em:< <http://www.w3.org/TR/voicexml20/>>. Acesso em: 26 jul. 2006

WORLD WIDE WEB CONSORTIUM. *Web Content Accessibility Guidelines 2.0*. [S.l.]:MIT;ERCIM; Keio, 2003. Disponível em: <<http://www.w3.org/TR/2003/WD-WCAG20-20030624/>>. Acesso em 11 ago. 2006

## REFERÊNCIAS CONSULTADAS:

*Jaws*(produzido pela enter-Joyce - empresa do grupo Freedom Scientific)

Disponível: <<http://www.freedomscientific.com>>. Acesso em: 12 out. 2004.

*Slimware Windows Bridge* (produzido pela Synthavoice)

Disponível: <<http://www.synthavoice.on.ca>>. Acesso em: 12 out. 2004.

*Virtual Vision* (produzido pela Micropower)

Disponível: <<http://www.micropower.com.br>>. Acesso em: 12 out. 2004.

Secretaria de Logística e Tecnologia da Informação - SLTI (artigos)

Disponível: <http://www.planejamento.gov.br> em 10/05/2005.

Modelo de Acessibilidade do Governo Eletrônico- e-MAG

Disponível:<<http://www.governoeletronico.gov.br/governoeletronico>>. Acesso em: 10 mai. 2005.

Recomendações para acessibilidade na *web* W3C em português - versão 1.0

Disponível: <<http://www.acessobrasil.org.br/index.php?itemid=41> >. Acesso em: 10 mai. 2005.

<http://www.acessobrasil.org.br/media/2/20030828-w3crevisto3.doc>

Forum VXML

Disponível: <<http://www.voicexml.org>>. Acesso em: 30 jun.2004.

SALT Forum

Disponível: <<http://www.saltforum.org>>. Acesso em: 04 set. 2004.

Nuance *Products*

Disponível: <<http://www.nuance.com>>. Acesso em: 18 out. 2004.

*Websphere Voice Server*

Disponível: <<http://www.ibm.com.br>>. Acesso em: 02 out. 2004.

*Oracle Collaboration Suite Voice*

Disponível: <<http://www.Oracle.com>>. Acesso em: 13 dez. 2004.

Microsoft SALT

Disponível: <<http://www.microsoft.com/speech>>. Acesso em: 13 dez. 2004.