

**INTRODUÇÃO****ASPECTOS  
PRELIMINARES****PROCESSAMENTO DE  
LINGUAGEM NATURAL****A LINGUAGEM NATURAL  
E SEUS COMPONENTES  
BÁSICOS****EXEMPLOS  
DE APLICAÇÕES****DETALHES DE  
IMPLEMENTAÇÃO****COMPARAÇÃO  
ENTRE  
SOFTWARES****FUTURO DA  
COMPREENSÃO  
DA FALA****GLOSSÁRIO,  
LINKS E  
REFERÊNCIAS****APLICATIVOS DE SINTETIZAÇÃO E RECONHECIMENTO DA VOZ**

Muito embora tudo o que foi dito até aqui com respeito ao reconhecimento da fala e da linguagem tenha se referido exclusivamente ao processo de síntese e produção da linguagem escrita, o que será exposto a seguir são aplicações que promovem o tratamento da linguagem natural de outra forma. Estes programas não se preocupam com tradução de textos, nem com solução de problemas de perguntas e respostas. O que eles fazem em sua maioria é promover uma interpretação do som de forma a reconhecer e diferenciar as palavras da linguagem natural, desempenhando determinadas ações conforme sua funcionalidade.

Neste caso, podemos caracterizar os aplicativos de fala computadorizada como pertencente a uma das duas categorias a seguir: aqueles que lêem texto para o usuário e aqueles que ouvem o que o usuário diz, seja para anotar um ditado ou para executar um comando. Em sua maioria, estes aplicativos podem ter seu desempenho melhorado através do treinamento. Isto acontece através da repetição das palavras por parte do usuário e fará com que o computador se acostume com seu estilo pessoal. Aplicações dotadas desta capacidade se tornam dependentes do orador.

Enormes são as dificuldades de compreensão da fala humana por um computador. A maioria destas dificuldades entretanto passam praticamente despercebidas pelas pessoas que se comunicam diariamente através da linguagem oral. Por exemplo, palavras de mesma pronúncia mas com significados diferentes (homônimas), como acento e assento, ou de palavras parecidas na grafia ou na pronúncia mas com significados diferentes (parônimas), como eminente e iminente, são facilmente distinguidas por pessoas através do contexto da comunicação, mas isto não é tão simples de ser representado em uma máquina.

Mesmo pessoas podem sentir grandes dificuldades às vezes com tais palavras. Por exemplo, durante o aprendizado de um novo idioma, palavras com pronúncia parecida se tornam irremediavelmente idênticas aos ouvidos mal treinados.

**LOGITECH AUDIOMAN**

Este aplicativo consiste em uma combinação de microfone e alto-falante de mão ou de mesa que é conectado a uma porta convencional de impressora.

Para o contexto presente, os aspectos mais interessantes do Audioman consiste em sua capacidade de converter textos em fala, bem como sua capacidade de receber textos ditados.

Um utilitário ainda permite a edição de diversas maneiras a gravação de voz. Os arquivos de som gravado do Audioman são do tipo WAV, compatíveis com o Windows.

Outro aplicativo do Audioman é o BESTpeech Read Out da Bekerley Speech Technologies, que converte texto em fala que inclui um dicionário de pronúncia, permitindo ao usuário fazer um ajuste fino do som das palavras.

A geração de voz promovida pelo BESTpeech se dá a partir de um conjunto de regras lógicas relacionadas ao som da língua falada, não sendo utilizado qualquer fragmento

derivado da fala humana real. Daí a voz monótona e com uma espécie de tonalidade mecânica e nasal produzida pelo programa.

Antes da aplicação do Audioman sobre um texto, este precisa ser selecionado em um editor de texto ou colocado na área de transferência do computador. Então, um normalizador expande as abreviações e coloca os números em sua forma por extenso.

Um dicionário de exceções trata das palavras que não se enquadram nas regras normais, após o que, as palavras são convertidas em seus equivalentes fonéticos.

Um conjunto de regras relacionadas ao ritmo da fala também tratam da ênfase e entonação das palavras, sendo que ajustes finos podem ser feitos sobre as regras de fonética.

Finalmente um gerador de voz entra em ação na geração da voz ouvida pelo usuário.

## **MICROSOFT WINDOWS SOUND SYSTEM**

O Microsoft Sound System é um conjunto de aplicativos de áudio fornecido com um fino microfone tubular, podendo ser adquirido em versões com ou sem placa de som.

O mais interessante de seus aplicativos para o contexto deste site é o Voice Pilot: um programa que permite associar qualquer comando vocal ou uma sequência de comandos de no computador.

Este aplicativo permite, por exemplo, que o usuário selecione um bloco de texto em um documento utilizando o método normal e depois dizer copy (copiar) ou paste (colar) para executar estes comandos. Com estes artifícios o usuário pode desencadear qualquer sequência de comandos do Windows.

Assim como outros aplicativos deste gênero, o Voice Pilot pode ser treinado pelo usuário para aprimorar sua capacidade de compreender seus comandos, analisando os padrões específicos de sua voz e salvando-os como gabaritos de áudio para serem consultados da próxima vez. O treinamento se dá como base na repetição das frases que o usuário julgar convenientes.

## **DRAGON SYSTEMS' DRAGONDICTATE FOR WINDOWS**

O Dragon Dictate é um aplicativo de ditado com extenso vocabulário (aproximadamente 60.000 palavras), reconhecendo a voz de qualquer pessoa, sendo que sua capacidade de reconhecimento melhora automaticamente com o tempo assim como o Voice Pilot.

O programa também se apresenta disponível em versões mais baratas com vocabulário reduzido de 5.000 e 30.000 palavras.

Quando surgem algumas dúvidas com respeito a alguma palavra pronunciada ele abre uma lista mostrando as hipóteses que lhe parecem mais viáveis, por ordem de probabilidade.

## **COMPUTADOR FALANTE**

Os laboratórios de pesquisa e desenvolvimento da British Telecom (BT), em Suffolk, no Reino Unido, trabalham com cabeças falantes há cerca de 10 anos. Durante esse período, suas criações já passaram da fase de imagens lentas e poucos convincentes, que demandavam enorme poder de computação da máquina, para fase atual, em que constituem retratos espantosamente realistas de pessoas reais, capazes de operar em qualquer computador pessoal de mesa.

As mais recentes cabeças falantes criadas pela BT conseguem adotar os traços faciais e as voz de qualquer pessoa. Elas sorriem, fazem cara feia, balançam a cabeça num

gesto de "sim" e até demonstram raiva, se aproximando cada vez mais do objetivo do pesquisador Andy Breen, que é gerar uma pessoa sintética capaz de passar ao usuário a sensação de que há alguém falando com ele.

O interesse da BT por pessoas sintéticas começou no início dos anos 80. Naquela época, o desafio era encontrar uma maneira de transmitir imagens de vídeo de boa qualidade pelas linhas telefônicas comuns.

Para reduzir a quantidade de informações a serem transmitidas, os pesquisadores tiveram a idéia de enviar apenas uma imagem, seguida por controles que modificavam elementos tais como movimentos dos lábios e as expressões faciais.

Embora tenha sido um sucesso parcial, o rosto resultante era pouco realista, seus movimentos muito abruptos e descontraídos.

A atual geração de cabeças falantes já avançou muito. Cada uma delas é baseada em uma armação tridimensional de arame, que é modificada para se adequar aos traços do rosto de uma pessoa específica.

Sobre a armação é superposta a imagem digital do rosto de uma pessoa, criando a ilusão de uma cabeça tridimensional. Os modelos das armações de arames simulam os movimentos de 13 músculos no rosto humano e, sua cabeça utiliza 19 diferentes modelos de boca para simular os formatos produzidos durante a fala normal.

Cada modelo incorpora modelo de lábios, dentes, maxilar e língua. Para isso, é selecionado um modelo ligeiramente diferente de armação de arame para cada som. Por meio de um teclado, o usuário registra texto que é convertido em palavras faladas pelos software Laureate, da BT, um sofisticado programa de conversão que converte o texto em discurso falado realista, com vários diferentes sotaques possíveis.

Os estágios intermediários entre eles são gerados e depois mostrados sucessivamente, dando impressão de movimento fluido.

O resultado final é semelhante a um filme sendo projetado a mais ou menos 10 quadros por segundo, e são gerados pelo computador em tempo real. Hoje eles estão procurando maneiras de criar mundos virtuais inteiros, nos quais as pessoas possam se encontrar e interagir.

"A imagem da pessoa é capitada por uma máquina fotográfica, traduzida em dados para computador e depois regenerada dentro de um mundo virtual", explica Breen.

Alguns exemplos iniciais dessa tecnologia já podem ser vistos na World Wide Web (WWW).

A tecnologia de cabeças falantes da BT vai levar este conceito ainda mais longe, possivelmente até mesmo permitindo que as pessoas usem seus próprios rostos.

Algumas aplicações futuras da tecnologia podem incluir secretárias pessoais eletrônicas que leiam as mensagens de e-mail que o usuário recebe e atendam a seus pedidos, além de guias virtuais que saíam à caça de informações e depois as apresentariam ao usuário sob forma oral.

## ELIZA

Eliza, um programa que alcançou alguma notoriedade, escrito há muito tempo, durante as primeiras experiências com linguagem e IA e que podia manter uma boa conversação com uma pessoa humana sobre uma certa variedade de tópicos.

Ele recebeu este nome devido à famosa personagem da peça de Shaw, "Pigmalião". Sua maneira de trabalhar era bastante simples, pois detectava a presença de algumas "palavras-chave" e, alterando os tempos verbais e extraíndo cláusulas das mensagens dos usuários, podia "dar a volta" nas sentenças e "dispará-las" de volta ao usuário.

A versão de maior sucesso deste programa podia conversar sobre quaisquer problemas que o usuário tivesse (ou imaginasse ter) e foi chamada de "Eliza Doctor".

Embora este programa fosse muito simples, tinha um grande vocabulário, que combinado com a maneira que os humanos "lêem significado" na conversa, o tornava muito convincente. Chegou a ser tão convincente que pacientes e doutores "reais"



começaram a usá-lo.

Na verdade, chegou a ser tão bom que muitos pacientes o preferiam a alguns médicos humanos!! O criador do programa e muitos outros especialistas em IA ficaram muito alarmados com a facilidade com que foi possível fazer um bom número de pessoas comuns e bem-educadas manterem longas e íntimas conversas com um programa completamente sem inteligência!

Os princípios usados pelo Eliza são muito fáceis de descrever. Inicialmente, varre a entrada em busca de algumas "palavras-chave", que, quando detectadas, disparam sempre a mesma ação, que consiste em ou devolver uma mensagem padrão, ou usar parte da entrada para construir uma mensagem. Por exemplo, se o usuário digitasse:

*Odeio sorvete*

O programa detectaria a palavra-chave "odeio" e responderia:

*Não é bom odiar.*

Observe que esta resposta era dada, independentemente do restante da frase. Se isto fosse tudo o que o Eliza fizesse, ficaria muito fácil descobrir que se tratava de um programa, pela quantidade limitada de respostas que daria. Entretanto, para introduzir alguma variação em suas respostas, o programa fazia uso de sentença de entrada.

Qualquer sentença que o usuário digitasse seria varrida em busca de certas palavras ou frases tais como "meu" é transformado em "seu" ou "você é" em "eu sou". A finalidade destas transformações simples é devolver ao usuário as sentenças que tenha introduzido, como se estas fossem geradas pelo programa. Por exemplo, recebendo a sentença:

*Você é um idiota.*

O computador devolveria:

*Sou um idiota.*

Possivelmente acompanhada de alguns sinais de exclamação ou de interrogação.

Estas duas técnicas - resposta a palavras-chaves e alteração de tempos verbais - acompanhadas de alguns outros truques especializados, podem produzir um programa que mantenha uma conversação razoável com você!

Existem à disposição programas Eliza em Basic que têm somente algumas páginas - outro testemunho do desejo dos humanos de darem um pouco de sua inteligência aos computadores! A moral desta história toda é que nem tudo que fala com alguém está entendendo o que é dito!

## **INTELLIGENT SPEECH ANALYSER™ (ISA) SOFTWARE** **(Analisador de Fala Inteligente)**

Os âmbitos principais de aplicação incluem: - Fonéticas - Phoniatics - Logopedics - Audiologia - Análise de Fala - Som Análises - Análise de Canções - Análise de Música - Análise de Instrumento de Música - Pesquisa choro Infantil - Pesquisa em Pulmão - Sons e Sons Cardíacos - Análise de Voz de Rádio - Editor de Som.

Todos os programas de análise foram escritos usando uma "máquina de linguas", porque desta forma, o ISA é muitas vezes mais rápido que usando um alto idioma nivelado. ISA é um software único no mundo.

O uso de ISA é muito simples. Todas as análises têm sua próprias janelas. Todas as funções do ISA são controladas pelo mouse. Todas as exibições podem ser escutadas. ISA-software está rodando em computadores Apple Macintosh. Macintosh é uma marca registrada da Apple Computer, Inc.

## **RUSSIAN TUTOR**

É um programa shareware para Windows 3.1 que ensina o básico do idioma Russo. Não é um programa tradutor de " frases" , mas sim um programa onde dá enfoques nas habilidades básicas que são necessárias para o estudo da língua.

Estas habilidades não são atingidas por memorização de palavras e frases, mas sim pelo estudo cuidadoso de sons, inflexão, e pronúncia diferenciada.

RussianTutor inclui algumas características em especial, como:

- \* Sete lições que cobrem os fundamentos do idioma russo, onde há vozes digitalizadas de um locutor russo nativo;
- \* Uso de Interface de Controle de Mídia (MCI) para todas as funções multimídia. Isto permite uso de qualquer hardware de multimídia MCI-compatível;
- \* Registro das vozes do estudante. O estudante pode comparar a própria voz com a das gravações.
- \* Interface vívida fácil-de-usar. Você precisará de ToolBook V1.53 (ou melhor) para rodar o Russian Tutor.

## **SISTEMA OPERACIONAL OS2**

As pessoas hoje em dia interagem de muitas formas com seus microcomputadores mas não estão usando o que há de melhor para esta interação, a própria fala.

O sistema operacional Os2 trabalha com reconhecimento de voz de forma muito eficiente. Podemos até dizer que é um dos melhores reconhecedores da fala. Por um método muito eficiente, neste sistema você é capaz de operar todos seus programas simplesmente com sua voz.

Para aumentar o desempenho e utilização deste recurso, o Os2 possui uma espécie de treinamento de usuário. Cada usuário pode definir seu gerenciador de voz ensinando o sistema a reconhecer seu modo de falar, incluindo tonalidade, ênfase, pronúncia e outros aspectos.

Este treinamento funciona mais ou menos na forma de ditado. O sistema pede para falar a palavra escrita na tela e após a execução, passa para outra e assim por diante. Desta forma, um usuário pode até definir sua própria voz para lhe servir de senha. Assim, se outra pessoa tentar utilizar o sistema, não funcionará uma vez que a voz é diferente da que foi configurada.

## **APOLLO 2**

O Apollo é um pequeno aparelho que se liga ao computador através da porta série. Possui um interruptor geral, alto-falante e controle de volume. Pode acomodar até sete idiomas, inclusive o Português.

O Apollo funciona em conjunto com um Software de Leitura de Ecrã ( por exemplo, o HAL ou o WINDOW BRIDGE). Este software permite explorar todas as possibilidades do Apollo: mudança da velocidade , do timbre e do volume da voz, entre outras.

Através do Apollo é de fundamental importância para usuários com deficiência visual navegarem na internet, pois este "traduz" as informações disponíveis na tela do computador para o usuário.

## **PARROT VOICE MATE**

O Parrot Voice Mate é outra tecnologia direcionada aos portadores de deficiência visual, trata-se de uma agenda pessoal com controle por voz.

As opções abaixo estão disponíveis na agenda e servem para facilitar a vida do usuário portador de deficiência:

- \* Possui marcação automática por voz de números de telefone.
- \* Calculadora controlada por voz, os resultados das operações são lidos aos usuários.
- \* Relógio com alarme e agenda de compromissos onde as informações (data, tipo e local de compromisso ou a hora se o usuário pedir, entre outros) são passados ao usuário por meio de uma leitura das mesmas.

O Parrot Voice Mate só é encontrado nos idiomas Inglês e Espanhol.

## VIA VOICE - IBM

O Via Voice é desenvolvido pela IBM para sistemas operacionais Windows.

Basicamente este software edita documentos profissionais e e-mail's através a voz natural do usuário, ou seja, o texto ditado pelo usuário é escrito na tela em um editor de textos qualquer.

Nas versões superiores é possível programá-lo para executar comandos específicos no desktop do usuário, sendo possível ativar programas apenas com comandos de voz.

O Via Voice pode ser acoplado a determinados browsers de modo que seja possível navegar na internet utilizando comandos ativados por voz;

Este software é muito bom e pode ser encontrado em português.

## TRADUTOR PALAVRA A PALAVRA - (ESCRITO EM PROLOG)

Este programa evita backtracking para zonas de erro quando há outras hipóteses;

```
% tradutor simples dict1(o,the).
```

```
dict1(gato,cat).
```

```
dict1(mia,miaus).
```

```
t1([],[]). t1([X1|L1],[X2|L2]) :- (dict1(X1,X2) -> true ; X2 = +X1 ), t1(L1,L2).
```

```
run(P,I) :- ex(P), t1(P,I). ex([o,gato,mia]).
```

```
ex([o,gato,estuda,ciencias,da,computacao]).
```

## TRATUDOR PALAVRA A PALAVRA - (ESCRITO EM PEARL)

Programa escrito em Perl com dicionario externo onde:

```
o=the -- gato=cat -- é=is -- um=a -- porco=pig
```

```
#!/usr/local/bin/perl
```

```
# traducao de palavras com consulta de um dicionario com validacao
```

```
# o dicionario e lido do ficheiro: dict
```

```
%dict=split(/[\n]/,`cat dict`);
```

```
while(<>) { s/(\w+)/if (defined $dict{$1}) { $dict{$1} } else {"+"$1"}/ge; print; }
```