

2. Jogos de aprimoramento da fala

Jogos de Voz

Os **Jogos de Voz** são exercícios de aprimoramento da coordenação fonoarticulatória de crianças e jovens surdos. Têm por objetivo aprimorar, de forma lúdica, o controle sensório-motor e destreza exigida para a realização dos sons, com mínimo esforço articulatório, comum aos exercícios fonoarticulatórios inseridos na área de atuação da fonoaudiologia.

Os **Jogos de Voz** foram criados dentro da comunidade acadêmica, no *Laboratório de Processamento Digital da Fala*, LPDF, na *Faculdade de Engenharia Elétrica e de Computação* da UNICAMP. A tese de doutorado, *Jogos Computadorizados Fonoarticulatórios para Crianças com Deficiência*, de Antonio Marcos de Lima Araújo, consistiu na ampliação e aprimoramento de um sistema anteriormente, feito durante os anos de 1994 a 1996.

As aplicações exigiram modelos matemáticos para proporcionar uma resposta visual automática do exercício fonoaudiológico executado pela criança. A variabilidade da voz, em especial de deficientes auditivos, exigiu a adaptação de algoritmos, que foram elaborados mediante um banco de dados criado por palavras isoladas e fonemas sustentados por crianças e adolescentes.

Foram desenvolvidos algoritmos que possibilitaram a realização dos aplicativos computacionais. Esses aplicativos foram desenvolvidos visando o aprimoramento do controle do fluxo respiratório, do controle da frequência de vibração das pregas vocais e do posicionamento articulatório. Além disso, esses aplicativos foram implementados sob a forma de jogos para motivar a criança com deficiência auditiva, usuária em potencial do presente trabalho.

Sua ênfase foi estabelecida atendendo requisitos específicos da engenharia elétrica e de computação. Por isso, é importante ressaltar que o enfoque dado ao desenvolvimento dos **Jogos de Voz** não compreendia a preocupação e o desenvolvimento de elementos visuais. Segundo o autor, que em março de 2003 concedeu uma entrevista para o presente trabalho, "os desenhos e as cores foram selecionados sem maiores critérios exceto que deveriam ser para utilização por crianças pequenas."

A criação dos jogos se deu através da experiência de atendimentos fonoaudiólogos e os movimentos das animações foram estabelecidos de acordo com os exercícios fonoaudiológicos aplicados durante a pesquisa.

O professor Antônio Marcos afirmou que a programação visual dos jogos não foi planejada e ressaltou que, após o desenvolvimento dos módulos, na fase de aplicação, percebeu a necessidade de modificar as imagens dos jogos, uma vez que uma das imagens do módulo de Caras "tem assustado algumas crianças muito pequenas."³⁹

Com os resultados da aplicação dos **Jogos de Voz** o autor verificou que houve um grande ganho na capacitação da fala em relação ao modelo anterior. Sobre a quantidade e qualidade, concluiu que "os alunos passaram a usar mais a emissão oral como forma de estabelecer contato com as outras pessoas"⁴⁰ e constatou que "a percentagem de vozes melhoradas foi de 80%. Os demais 20% não apresentaram melhorias na voz, não havendo em nenhum caso piora da voz."⁴¹

Ao desenvolver jogos de computador que visam representar visualmente um fenômeno natural através de um sintetizador de voz, além da capacidade técnica, o que os diferencia dos outros jogos para a aquisição da fala de crianças surdas é sua interação com a criança e a forma de transmissão da informação. As crianças interagem com o jogo através de um microfone, que sintetiza sua voz e a transforma em imagens que representem as características dessa voz. As possibilidades de representações gráficas são as mais variadas, assim como as suas dinâmicas.

O trabalho desenvolvido por Araújo se limitou ao tratamento da *fala* partindo do pressuposto que essa é essencial para a construção da linguagem, e é utilizada na interação com o outro.⁴² Citando Albano, Araújo descreve que a presença da fala colabora na redescoberta de processos de relação com a própria linguagem, dando uma gramática aos recursos já explorados ludicamente.⁴³

Os **Jogos de Voz** são um conjunto de quatorze módulos que consistem nos jogos de controle de energia, exercícios de respiração, exercícios das fricativas, frequência fundamental e vogais do português. O meio de relação entre jogador e o jogo é a captação da voz do jogador, isto é, a partir da execução da fala são dados os comando e as respostas visuais. Os jogos foram realizados em C Builder, e os requisitos mínimos da

³⁹ A entrevista concedida em março de 2003 pode ser vista na íntegra na seção anexa do trabalho, p. **Erro! Indicador não definido.**

⁴⁰ Araújo, A..M.L. *Jogos Computadorizados Fonoarticulatórios para Crianças com Deficiência Auditiva* p.101

⁴¹ ARAÚJO, A..M.L. Op. Cit. p.102

⁴² ARAÚJO, A..M.L. Op. Cit p.9

⁴³ ARAÚJO, A..M.L. Op. Cit p.10

plataforma computacional para operação são: um computador tipo PC, com relógio de frequência maior ou superior a 100MHz, dotados com placa de som e microfone, operando sob o sistema operacional Windows.

Os quatorze módulos dos jogos de voz são os módulos de **Caras**, **Espaço**, **Futebol**, **Fricativas FV**, **Fricativas XJ**, **Fricativas SZ**, **Fricativas FSX**, **Fricativas VZJ**, **Pássaros**, **Helicóptero**, **Tiro ao Alvo**, **Vogais da Aranha**, **Vogais do Espaço** e **Vogais do Macaco** e serão descritos abaixo.

O módulo de **Caras**, assim como **Espaço** e **Futebol**, visa controlar a energia pelo aprimoramento do controle de intensidade da fala. Modula a voz, estabelecida por faixas de intensidade e captadas através de um microfone, exibindo para a criança três figuras (Figura 1). Cada figura representa faixas determinadas de intensidade sonora: baixo, ideal e alto. Quando a criança fala na frequência ideal estabelecida pelo fonoaudiólogo, a figura A se transforma na figura B. O objetivo é manter a figura B aprendendo a controlar os níveis de intensidade. O autor relaciona a imagem B com a faixa ideal determinada pelo fonoaudiólogo nas configurações iniciais e denomina a figura como estímulo positivo. Quando a criança passa o limite estipulado, a figura B se transforma na figura C. As figuras A e C são consideradas estímulos negativos.

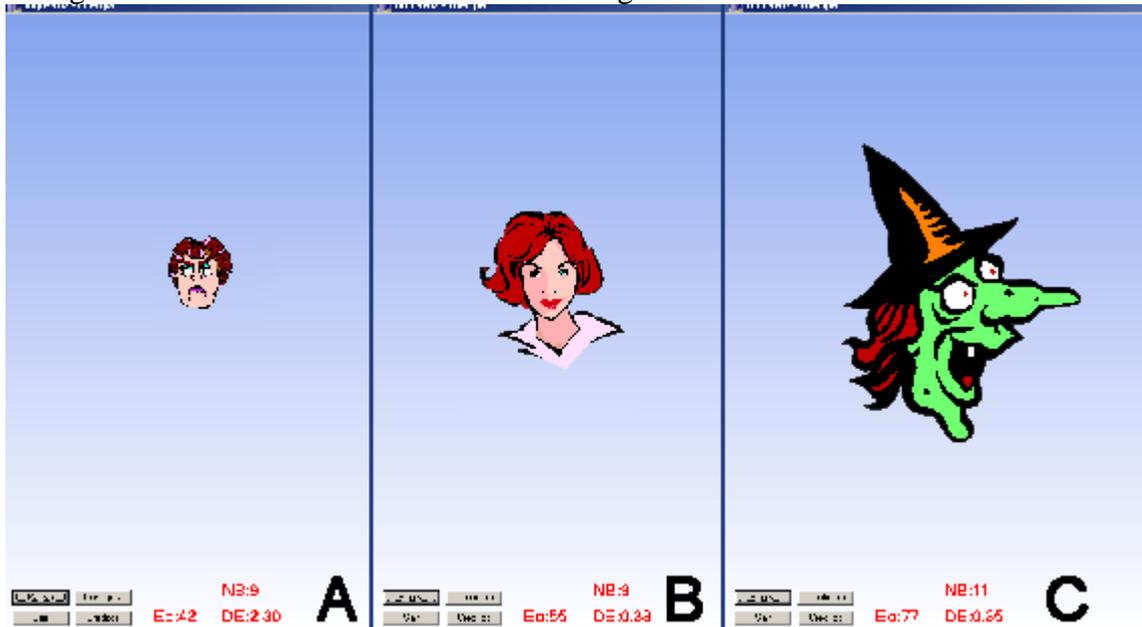


Figura 1 - Jogos de Voz - Módulo de energia - Caras

No jogo do **Espaço** (Figura 2), o objetivo é apanhar os astronautas com a nave espacial que é controlada pelo nível de intensidade da fala. A nave move da esquerda para direita enquanto os planetas e os astronautas da direita para a esquerda. A nave deve recolher os astronautas, aumentando os pontos e se desviar dos planetas, evitando perder os pontos. O nível de intensidade controla a altura da nave espacial na tela, isto é, a

nave é controlada aumentando e diminuindo a intensidade da VOZ.



Figura 2 - Jogos de Voz Módulo de energia - Espaço

No jogo **Futebol** (Figura 3), a criança modula a intensidade da fala acertando o gol. Se a intensidade estiver inferior a estabelecida, a bola é jogada para baixo do gol (em relação à tela). Se a intensidade estiver acima, a bola também irá para cima. No jogo de **Futebol** é preciso manter o controle de intensidade acertando o gol.



Figura 3 - Jogos de Voz Módulo de energia - Futebol

Nos cinco módulos de **Fricativas** (Figura 4) a criança deve desenvolver as consoantes /f/, /s/, /x/, /v/, /z/ e /j/. Cada fricativa é associada a uma coluna horizontal, como uma barreira que será aberta para o pássaro atravessar apenas se a

criança produza o som suficiente para abrir a barreira ou o pássaro baterá na parede e voltará.



Figura 4 - Jogos de Voz – Fricativas

O módulo do **Pássaro** (Figura 5) tem a finalidade de coordenar a respiração e aumentar a resistência cardio-respiratória do jogador. O objetivo do jogo é colocar o objeto na porção superior da tela, exigindo um aumento da duração da expiração e uma diminuição do período de inspiração, simulando a respiração para a fala e proporcionando o aprimoramento do controle da fonte. O controle da respiração é um dos pontos primordiais para o aprendizado da fala. Os sons na maioria das línguas são produzidos durante a expiração. As crianças surdas apresentam, em geral, um tempo de fonação menor do que as crianças ouvintes.



Figura 5 - Jogos de Voz - Módulo de Respiração – Jogo do Pássaro

No jogo do **Helicóptero** (Figura 6), a frequência fundamental é obtida de acordo com o percurso do helicóptero: ele sobe enquanto estiver na frequência adequada e desce quando a locução produzida estiver acima ou abaixo da

estabelecida nas configurações. O objeto subirá quando a frequência da locução estiver na faixa especificada e descerá quando a locução produzida apresentar frequência acima ou abaixo do valor estabelecido. A frequência fundamental ou frequência de vibração das pregas vocais é uma característica individual, cujos valores médios dependem da idade e do sexo. A vibração das pregas vocais ocorre em mais de 83% dos fonemas utilizados na língua portuguesa.

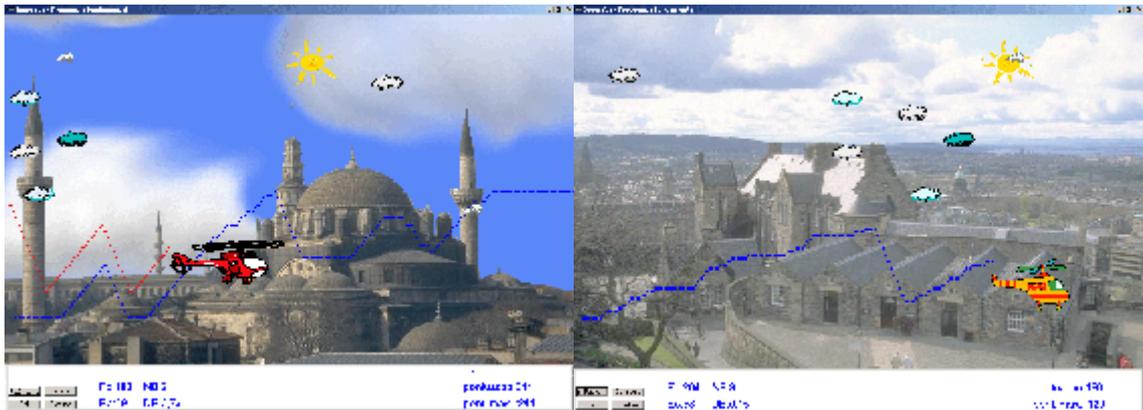


Figura 6 - Jogos de Voz – Módulo de frequência fundamental – Helicóptero

Outro jogo de frequência fundamental é o **Tiro ao Alvo** (Figura 7). A criança modula a voz fazendo com que a árvore acerte as frutas no alvo de acordo com a faixa de frequência estabelecida. O agudo faz com que as frutas passem por cima do alvo. Quando a criança impõe a voz mais grave, ela acerta o alvo.



Figura 7 - Jogos de Voz – Tiro ao alvo

Os módulos de exercício das **Vogais do Macaco** (Figura 8), **do Espaço e da Aranha** consistem em incentivar a criança a falar a vogal assinalada pelo software. Quando a criança fala

corretamente a vogal /o/, por exemplo, assinalado com uma banana, o macaco busca essa banana.

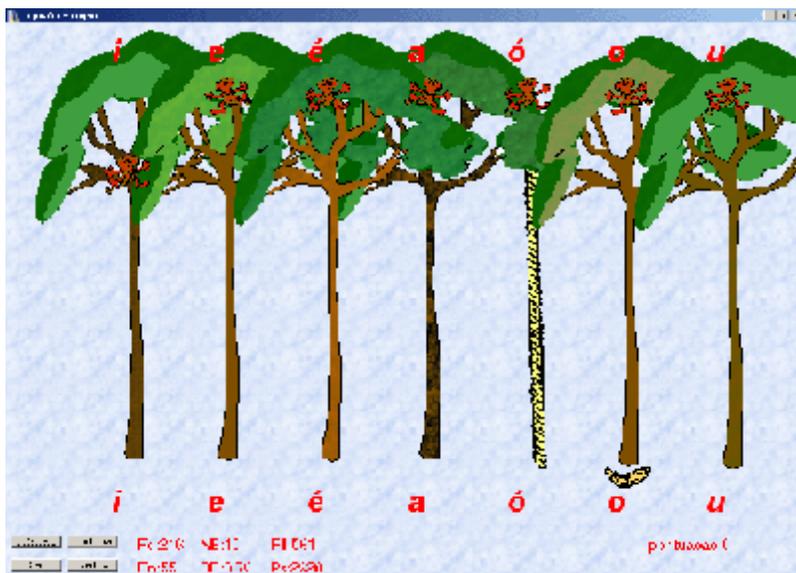


Figura 8 - Jogos de Voz – Módulo de Vogais

Todos os módulos dos Jogos de Voz compreendem sua representação visual pela *retroalimentação (feedback) visual*. Isto é, representam visualmente as características da fala no instante que é falado. Através da modulação da voz, a imagem também é modulada, tentando substituir uma sensação inexistente ou pouco desenvolvida nas crianças surdas: a *retroalimentação auditiva*.

Conceito dos jogos: em tempo real

Os **Jogos de Voz** utilizam como princípio o conceito da retroalimentação (*feedback*) visual para substituir a natural retroalimentação auditiva, ausente nos surdos, utilizada na modulação da voz. Por isso, para compreender a estrutura dos jogos é necessário entender o conceito da retroalimentação e suas origens.

O termo *retroalimentação*, ou *realimentação*, do inglês *feedback*, é um conceito muito utilizado em sistemas eletrônicos, também aplicado em outras áreas do conhecimento, como, entre outras, a biologia, a matemática, a comunicação. Tem o objetivo de controlar e auto-regulamentar sistemas através de mensagens de retorno. É um método de verificação que dá ferramentas para a análise e compreensão de sistemas.⁴⁴ Fundamentado na *Teoria Geral dos Sistemas*, desenvolvida nos anos 40 pelo biólogo Ludwig von Bertalanffy (1901-1972), que defende um modelo científico explicativo do comportamento de um organismo vivo, progênie da Termodinâmica, o termo *feedback* foi desenvolvido e aplicado por Norbert Wiener (1894-1964), juntamente com os conceitos da *cibernética*.

⁴⁴ Enciclopédia Microsoft Encarta 2001.

Derivada da palavra grega *Kubernetes*, (significado para "piloto" ou "timoneiro" e derivação da palavra "governador"), a *Cibernética* é ciência que trata dos sistemas de comunicação e controle dos organismos vivos, das máquinas e das organizações. O termo foi aplicado por Wiener, em 1948, para representar a *teoria dos mecanismos de controle*. A *cibernética* investiga técnicas pelas quais a informação se transforma na ação desejada. Essa ciência surgiu dos problemas levantados durante a II Guerra Mundial quando do desenvolvimento dos chamados cérebros eletrônicos e dos mecanismos de controle automático para equipamentos militares, como, por exemplo, os visores de bombardeio.⁴⁵ Anteriormente, o termo havia sido utilizado por Ampère, nos primórdios do Século XIX, com referência à ciência política⁴⁶.

O conceito designado como *feedback* é mais antigo que a *cibernética*, antecede o princípio de conduzir um navio (de onde foi extraído seu nome). Buscando a origem do conceito, podemos verificar que Hipócrates (460?-377 a.C.), nos textos que compõem o *Corpus Hippocraticum*, escritos entre 420 a 350 a.C., considerou o corpo humano como um sistema influenciado pelo meio ambiente, na qual a saúde é o equilíbrio de quatro humores (sangue, bÍlis amarela, bÍlis preta e fleuma) que, quando em desarmonia, provocam algumas mensagens como vômito, tosse, suor excessivo, urina escura, entre outros. O papel do médico, para Hipócrates, seria interpretar estas mensagens, descobrindo como se dava esse desequilíbrio, para assim, diagnosticar o doente⁴⁷. Esta é uma das mais genuínas idéias de *feedback* e mensagem.

Wiener defende a tese de que a sociedade só pode ser compreendida através de um estudo das mensagens e das facilidades de comunicação de que disponha⁴⁸, onde a *cibernética* tem o propósito de desenvolver linguagem e técnicas que capacitem o controle e o repertório de técnica e idéias adequadas para classificar as certas manifestações e respostas. Segundo o autor,

o homem está imerso num mundo ao qual percebe pelos órgãos dos sentidos. A informação que recebe é coordenada por meio de seu cérebro e sistema nervoso até, após o devido processo de armazenagem, colação e seleção, emergir através dos órgãos motores geralmente os músculos. Esses, por sua vez, agem sobre o mundo exterior e reagem, outrossim, sobre o sistema nervoso central por vias de órgãos receptores, tais como os

⁴⁵ Enciclopédia Microsoft Encarta 2001.

⁴⁶ WIENER, N. *Cibernética e Sociedade: O Uso Humano dos Seres Humanos*. p. 15

⁴⁷ Enciclopédia Microsoft Encarta 2001.

⁴⁸ WIENER, N. Op. Cit. p. 16

órgãos terminais de cinestesia; e a informação recebida pelos órgãos cinestésicos se combina com o cabedal de informação já acumulada para influenciar as futuras ações.

Pela cibernética, a retroalimentação (*feedback*) é o controle da máquina, baseado no desempenho efetivo (e não no desempenho esperado), que tem a função de detectar e monitorar elementos⁴⁹.

Wiener afirma que a totalidade da vida social humana se centra na linguagem e, quando da ausência de seu aprendizado no devido tempo, o conseqüente malogro do aspecto social do indivíduo.⁵⁰

Em síntese, propõe que a retroalimentação é um controle dado pela reintrodução dos resultados do próprio desempenho; e que a aprendizagem é concebida mediante a capacidade de modificar o desempenho do sistema pela informação ou mensagem estabelecida.⁵¹

Segundo Mautner⁵²,

motora ou mental, a aprendizagem se dá sempre no campo da tentativa e erros corrigidos por '*feedback*'. Se os antropólogos tiverem razão, isso é válido para todos os povos em todas as épocas. É errando, percebendo o erro e corrigindo-o que se chega ao acerto. É escorregando e caindo que se aprende a andar.

O mundo das telecomunicações e da informática está propiciando novas maneiras de pensamento e relacionamento entre os homens e o meio. "Escrita, leitura, visão, audição, criação, aprendizagem são capturados por uma informática cada vez mais avançada. [...] Emerge, neste final do séc XX, um *conhecimento por simulação*..."⁵³

Estudos que envolvam o processamento da fala, sua decodificação em sinais e transporte, vêm se aprimorando ao longo do tempo porque se apropriam de novas tecnologias desenvolvidas nas áreas de comunicação e informação.

Por volta de 1949, o interesse pela utilização de máquinas que auxiliassem pessoas com algum tipo de deficiência, fez com que Norbert Wiener utilizasse o conhecimento do telefone, que ele considerava ser a mais desenvolvida das técnicas de comunicação, porém colocando um enfoque na capacidade de condução do sentido tacto⁵⁴. Wiener defendia, a partir da análise de outros trabalhos, como o

⁴⁹ WIENER, N. Op. Cit. p. 24

⁵⁰ WIENER, N. Op. Cit. p. 84

⁵¹ WIENER, N. Op. Cit. p. 61

⁵² MAUTNER, A.V. *A calçada, o campinho e o computador*

⁵³ LÉVY, P. *As Tecnologias da Inteligência. O Futuro do Pensamento na Era da Informática.* p.7

⁵⁴ WIENER, N. Op. cit. p. 165

desenvolvido nos anos 40 pelos *Laboratórios da Companhia Bell*, conhecido como *Visible Speech*, e seu antecessor *Vocoder*, que a fala visível tinha duas desvantagens: "não parecia ser fácil de produzir em forma portátil, e fazia exigências excessivas ao sentido da visão, o que é mais importante para uma pessoa surda do que para o resto de nós."⁵⁵

Podemos notar que, atualmente, estas desvantagens não são mais empecilhos para as pesquisas em retroalimentação. Primeiro porque os sistemas estão a cada dia mais portáteis e velozes, segundo porque a comunidade surda, assim com qualquer sociedade em geral, está mais condicionada à comunicação visual. Constatamos que a utilização de mensagens visuais é mais uma forma de auxílio na aquisição da fala e da linguagem de crianças surdas.

Wiener distingue, de forma sucinta, "três estágios de linguagem, e duas de traduções intermediárias, entre o mundo exterior e o recebimento subjetivo de informação", ao tentar interpretar a fala através de outro sentido.

O primeiro estágio consiste nos símbolos acústicos tomados fisicamente como vibrações no ar; o segundo, ou estágio fonético, consiste nos vários fenômenos do ouvido interno e da parte conexa do sistema nervoso; o terceiro, ou estado semântico, representa a transferência desses símbolos para uma experiência de significação.⁵⁶

Wiener afirma a ausência do segundo estágio nas pessoas surdas, porém, esta conexão comum ao ouvido pode ser substituída por outro sentido como o tato ou a visão, a partir de um sistema artificial e não o sistema físico-nervoso natural. Contudo, a translação do novo segundo estágio para o terceiro representa um novo sistema de hábitos e respostas que é comparado à aprendizagem de uma nova função como a de dirigir um carro. A primeira tradução, da primeira fase para a segunda, é facilmente controlada, enquanto a segunda tradução, da segunda para a terceira fase, mostra ser puramente o processo de aprendizagem.⁵⁷

O neurologista Oliver Sacks⁵⁸ exemplifica a importância fundamental do *feedback* sensorial (os sentidos conhecidos, acrescidos da propriocepção)⁵⁹ para que o homem reconheça

⁵⁵ Wiener, N. Op. cit. p.166

⁵⁶ WIENER, N. Op. cit. p. 166

⁵⁷ WIENER, N. Op. cit. p. 166-7

⁵⁸ SACKS, O. *O Homem que Confundi sua Mulher com um Chapéu*. p.63-65

⁵⁹ Os cinco sentidos humanos são distintos pela visão, audição, tato, olfato e paladar. Os receptores em jogo são o olho, ouvido, planta dos dedos, nariz e língua que constituem os órgãos sensitivos. (Larousse Medical – sens). Um outro sentido, chamado por Sacks de sexto sentido, é a propriocepção. Batizada na década de 1890, a propriocepção é a percepção que o homem tem do próprio corpo humano, isto é, uma

seu meio ambiente. Ao considerar o senso do corpo dado pela união de três sentidos: a visão, a audição e a propriocepção, afirma que, na ausência de um deles, os outros poderão compensá-lo ou substituí-lo, em certa medida. No caso "a mulher desencarnada"⁶⁰, a paciente se apropriou do *feedback* auditivo e visual para suprir o *feedback* da propriocepção. Um *feedback* inconsciente cedeu lugar a um consciente.

Segundo Sacks, o *feedback* auditivo "é secundário e tem pouquíssima importância para a fala. Pois a modulação da fala normalmente é proprioceptiva, governada por impulsos recebidos de todos os nossos órgãos vocais."⁶¹ Por isso surdos congênitos podem adquirir uma fala praticamente perfeita.

A maioria dos surdos tem o aparelho fonador idêntico ao dos ouvintes, isto é, não são mudos, e poderiam falar da mesma forma que ouvintes. O que lhes falta é a capacidade de ouvirem a própria fala e, portanto modular e regular a voz pelo ouvido.⁶²

As pessoas com surdez *pré-linguística*⁶³, "não dispõem de imagem auditiva, não têm idéia alguma de como é realmente o som da fala"⁶⁴, não existe correspondência entre som e significado. Para aprender a monitorar a fala, os surdos precisam buscar outros sentidos, como a visão, o tato, a cinestesia e o senso de vibração, para representar o papel da audição.⁶⁵

A monitoração da fala pelo ouvido dos ouvintes é, portanto, o que podemos chamar de retroalimentação (*feedback*) auditiva existente na pronúncia da palavra. Quando a criança ouve o que é dito por ela e pelos outros ela pode modular sua voz através da percepção auditiva. Isso é uma maneira natural de aquisição das palavras. Crianças surdas não têm possibilidades de analisar a voz e a entonação do pronunciamento de palavras pelo mesmo órgão que os ouvintes, o que faz da percepção visual uma das soluções compensatórias.

Segundo Pritch⁶⁶, o sentido visual tem capacidade suficiente para controlar os órgãos audíveis. O conceito básico da retroalimentação visual pode ser observado no diagrama da

percepção, que associada às outras, dá ao homem a noção do espaço e equilíbrio.

⁶⁰ SACKS, O. Op. cit.

⁶¹ SACKS, O. Op. cit. p.65

⁶² SACKS, O. *Vendo Vozes*. p. 38

⁶³ Sacks considera pessoas com surdez *pré-linguística* aquelas que nasceram surdas ou ficaram antes de adquirirem uma linguagem. Diferente das pessoas com surdez *pós-linguística*, que, apesar de também não monitorarem a fala pelo ouvido, podem utilizar da lembrança para executar a fala de forma regular.

⁶⁴ SACKS, O. Op. Cit. p. 39

⁶⁵ SACKS, O. Op. Cit. p. 38

⁶⁶ PRITCH, M. *Visual Communication Aid for Deaf People*. p.1

Figura 9. Um ciclo na qual parte da fala da criança surda (D), que é gravada no computador através de um microfone (A), decodificada e transformada em imagem e exibida imediatamente na tela do monitor do computador, possibilitando ao usuário, que recebe esta informação visual (B), controlar e corrigir a pronúncia (C) e, assim, utilizando a representação visual da fala.

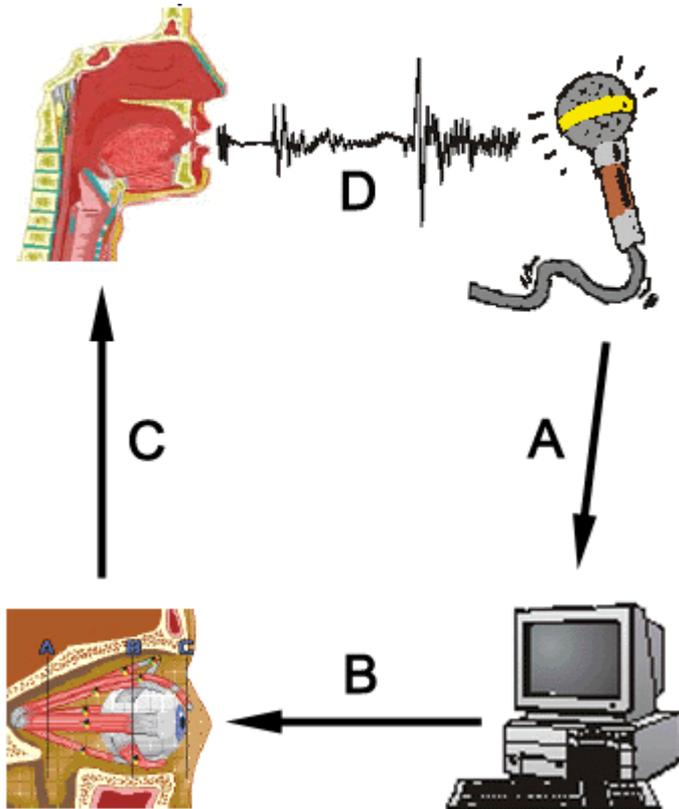


Figura 9 - representação do sistema de retroalimentação visual.

O *feedback* visual encontrado nos **Jogos de Voz** pode ser pensado à luz dos conceitos da *cibernética*. Utilizaremos, no decorrer da dissertação, o termo *feedback* para nomear o que chamamos aqui de *retroalimentação* visual.

Igual, mas diferente

Não são recentes as tentativas de utilizar imagens para substituir a uma ausência do sentido auditivo e, como pode ser visto, a utilização de máquinas para isso surge concomitante às tecnologias. Os Laboratórios Bell são um exemplo disso, as tecnologias de decodificação do som em sinais contribuíram para o advento do telefone.

Em 1874, Alexander Graham Bell idealizou um "display visual", que apresentava instantaneamente as ondas sonoras do usuário. O "*Visible Speech Translator*" foi criado em 1944 no Bell Telephone Lab e consistia na retroalimentação visual das

ondas sonoras em uma tela parecida com a de uma televisão⁶⁷. Depois disso, o desenvolvimento de programas de retroalimentação visual em crianças surdas se cresceu no final dos anos 50, início dos 60. A entrada do computador nesse processo possibilitou maior flexibilidade na programação e visualização da fala. Existe no mercado um pequeno número de softwares de treinamento da fala baseados em computador, comercializado.

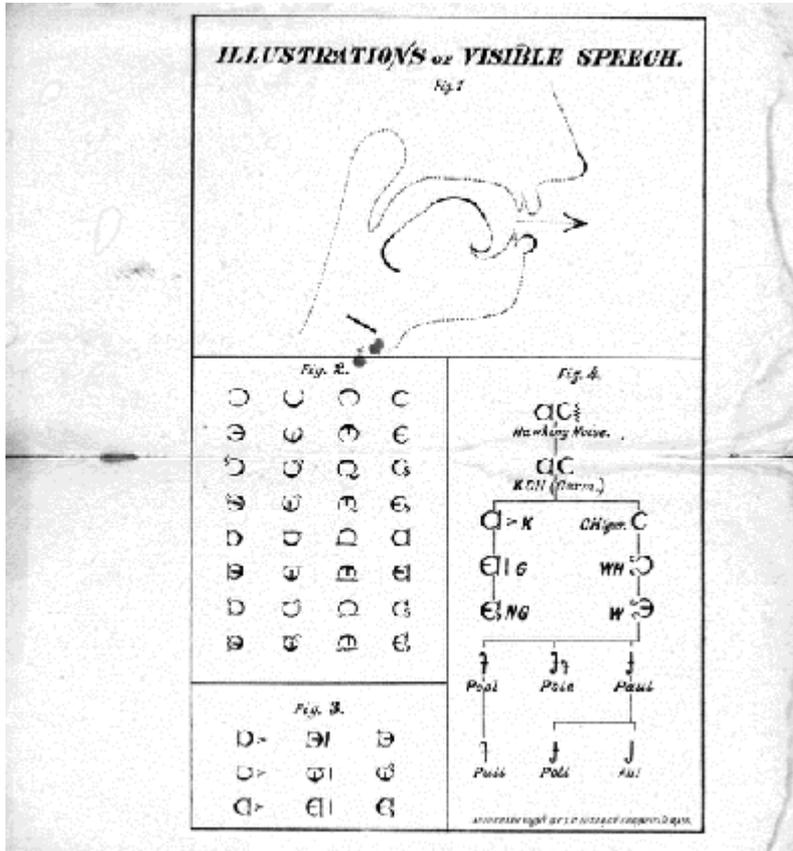


Figura 10 - estudos iniciais de representação gráfica do "visible speech" desenvolvido posteriormente pelos Laboratórios Bell

Existem estudos significativos que utilizam a retroalimentação visual, uma vez que a tecnologia advém das pesquisas em telecomunicações.

Muitos dispositivos que usam a informação visual como alternativa para auxílio da língua falada e escrita de pessoas surdas.

Os primeiros sistemas o comportamento era indicado por uma lâmpada especial ou pela deflexão de uma agulha do medidor. Alguns sistemas usaram osciloscópios revelar informações mais detalhadas dos padrões acústicos da fala. Apesar do fato de que muitos desses dispositivos automáticos terem relatado a melhora da fala de algumas crianças, o uso deles era limitado. Isto provavelmente devido ao fato que o gabarito visual não possibilitava uma

⁶⁷ PELLEGRINO, C.N. *Estudo em Imagens Falantes* p.13

compreensão fácil ou natural desses dispositivos de treinamento.

Uma contribuição significativa foi feita por Nickerson & por Stevens, em 1976, quando desenvolveram o primeiro sistema por computador do treinamento da fala. Desde então, o desenvolvimento técnico do computador forneceu softwares de treinamento da fala mais avançados. A maioria desses sistemas contém um microfone, um amplificador e um alto-falante conectado que permite a entrada do som no computador (input), armazena e analisa o discurso e então o indica rodando uma imagem na tela do computador.⁶⁸

Iguais, mas diferentes, porque não é possível utilizar softwares estrangeiros (levando em consideração que foram feitos na língua de origem) sem quase transformá-los em outros. A especificidade de software para aquisição da fala é muito grande, devido às diferenças estruturais e lingüísticas de cada língua. Independentemente das pesquisas e *know-how* internacional que podem ser obtidos, softwares de aprimoramento de fala da língua portuguesa falada no Brasil necessitam de investimento interno.

Alguns exemplos do funcionamento dos softwares estrangeiros estão descritos nos anexos do trabalho. Um deles, entretanto, tem seu conceito muito semelhante aos **Jogos de Voz**. O *Skill builder* foi desenvolvido pela fábrica de softwares *Doctor Speech* para crianças de língua inglesa. Os criadores do programa afirmam que o caráter lúdico, as imagens e as cores, além da interatividade dos jogos tornam a terapia mais dinâmica e alegre. O software propõe ao usuário atingir metas durante um determinado tempo. Utiliza o computador pessoal (PC), microfone, placa e caixa de som. O software analisa parâmetros como potência, altura, voz, tempo de fonação com retroalimentação e performance automática.

O *módulo de potência* ajuda na modulação da voz e tem o mesmo objetivo do módulo de **Caras** dos **Jogos de Voz**. Como exemplo, na Figura 11, a criança pronuncia a vogal /a/ na frente do microfone fazendo com que o barquinho navegue sem colidir com as pedras. A representação gráfica de Hz⁶⁹ por tempo tem a curva similar ao movimento do barquinho (Figura 12).

⁶⁸ ÖSTER, A., HOUSEN, D., PROTOPAPAS, H, HARZIS, A. *Presentation of a new EU project for speech therapy: OLP (Ortho-Logo-Paedia)*. p.3 (minha tradução)

⁶⁹ Símbolo para Hertz: Unidade de medida de freqüência de um fenômeno periódico



Figura 11 - Skill Builder – módulo de potência

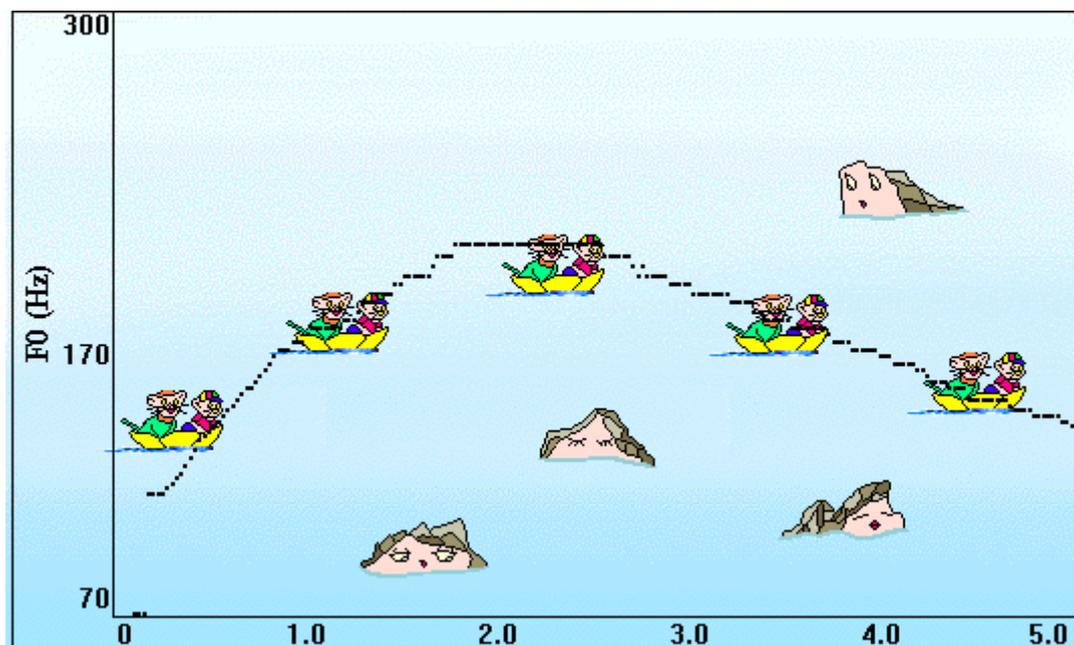


Figura 12 - representação vetorial da curva de potência sobreposta a imagem

O *módulo de energia* ajuda a criança a estabelecer e dominar a altura da fala. Uma forma ensinar o controle da altura da voz é através do controle correto da respiração: um bom posicionamento corporal ajuda a controlar os volumes da fala. Na Figura 13, aumentando o volume da fala, o bombeiro sobe a escada para resgatar o gatinho em cima da árvore. O volume de alcance pode ser estabelecido pelo profissional que dá suporte ao atendimento.



Figura 13 - Skill Builder – módulo de energia – representação visual

Os equipamentos computacionais de *feedback* visual de apoio à terapia de fonoaudiologia podem ser classificados em três categorias de softwares. A primeira inclui todos os produtos sem interatividade entre o usuário e o sistema. Um exemplo é o *HeCom Sprachlaute*, desenvolvido na língua alemã, que mostra as posições corretas da língua e da boca em cada fonema, como podem ser vistos na Figura 14. Esta categoria é muito parecida com a forma tradicional de terapia, quando utiliza espelhos ou pela imitação dos movimentos executados pelo fonoaudiólogo.

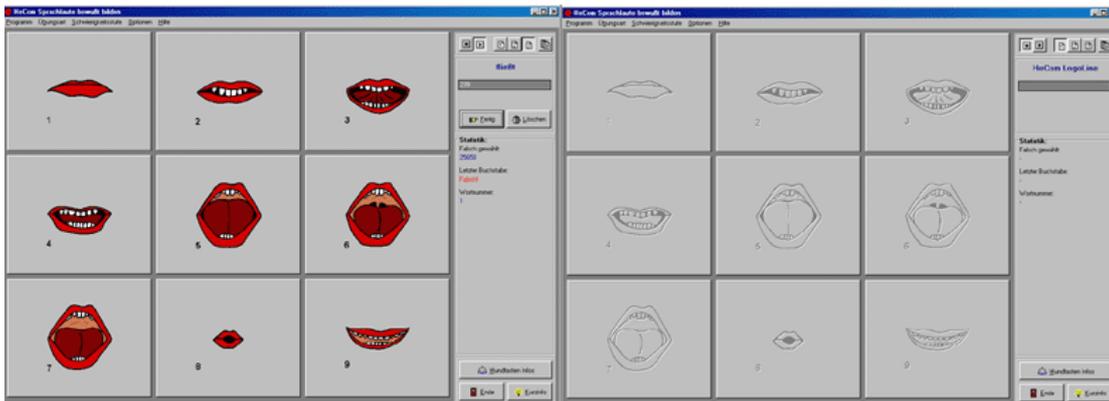


Figura 14 - HeCom Sprachlaute – parecido com a fonoaudiologia convencional

A segunda categoria inclui todos os produtos que permitem interação direta com o sistema. São jogos como os desenvolvidos pela IBM, chamado *Speech Viewer* e os **Jogos de Voz**. Utilizam animações, desenhos, gráficos e espectrogramas como síntese do som. A grande maioria é desenvolvida para o público infantil e propicia a estimulação precoce da fala em crianças surdas. A mostra uma das imagens utilizadas no *Speech Viewer III*.

A criança produz efeito sonoro constante até o ursinho, que se movimenta de acordo com locução, chegar no bloco verde (Figura 15). O *Speech Viewer* tem capacidade de gravar a própria voz e a criança recebe a retroalimentação visual. Os pequenos jogos demonstram na tela como controlar a volume, a entonação, a frequência, a potência, a respiração, as vogais, as fricativas e os fonemas.

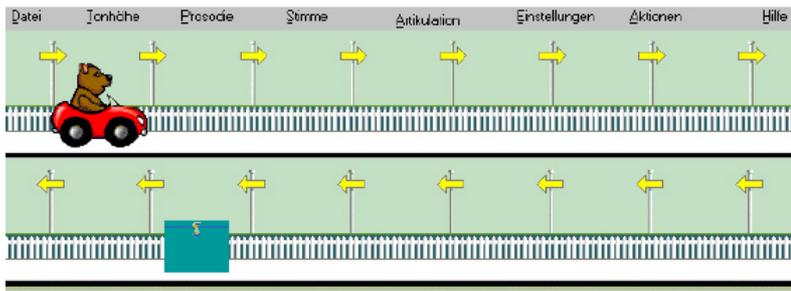


Figura 15 - Speech Viewer III

A terceira categoria inclui softwares em que os sinais de voz podem ser manipulados. Um exemplo é o programa *Sprechtrainer* da HB, desenvolvido na Alemanha, que mostra a posição correta dos lábios e região vocal e compara o sinal gravado com o sinal de referência. Consiste na exibição de várias telas e depende de uma compreensão maior do usuário, comparado às anteriores. Seu uso é mais comum no auxílio de jovens e adultos. O programa se expressa por julgamentos como excelente, bom ou ruim⁷⁰, como pode ser visto na Figura 16.

⁷⁰ PRITCH, M. *Visual Communication Aid for Deaf People*. p.4

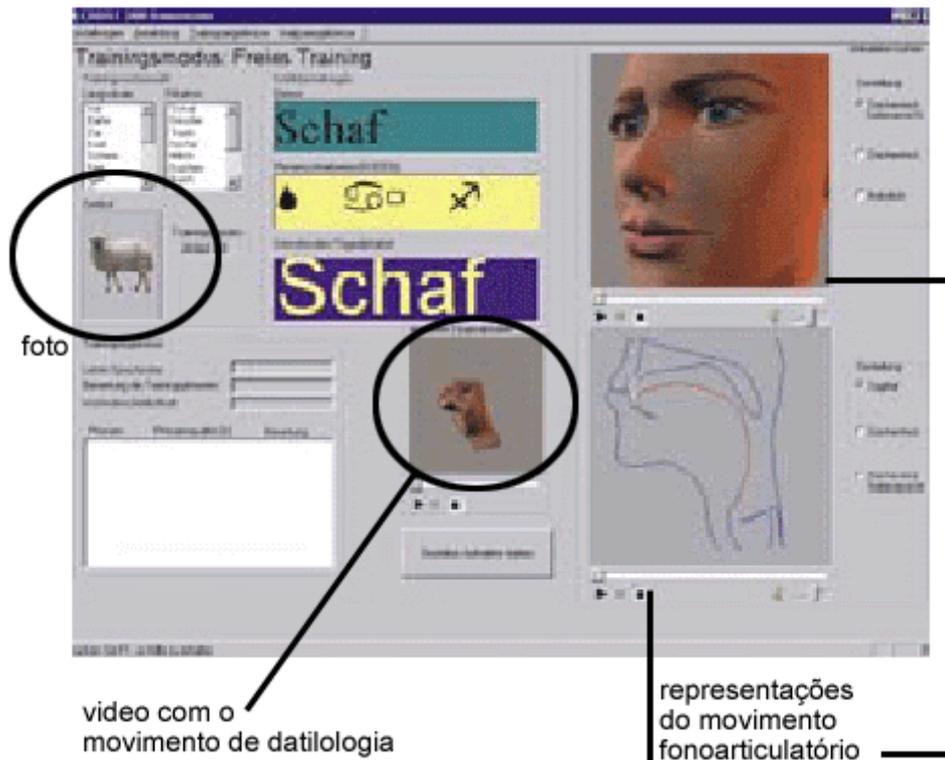


Figura 16 - Spreachtrainer

No *Sprechtrainer* aparecem para uma palavra como *schaft* (ovelha) várias representações: uma foto do animal, o sinal em datilologia, o símbolo do zodíaco, o movimento fonoarticulatório da palavra executado por um rosto virtual, a vista ortogonal do movimento fonoarticulatório, o signo textual na família de fonte com ou sem serifa e o espectrograma da palavra.

Segundo Öster⁷¹, existem pontos de grande relevância nas características do software de *feedback* visual para seu bom funcionamento. São eles:

- as instruções e os manuais pedagógicos devem ser elaborados de forma clara, possibilitando o uso de diferentes grupos de crianças;
- a retroalimentação visual da voz e dos movimentos deve aparecer de forma instantânea, sem nenhum atraso;
- o apoio deve ser aceito tanto pelo fonoaudiólogo quanto pela a criança, isto é, o software deve ser atrativo, interessante, de fácil compreensão, fácil manuseio e motivador;
- o ambiente visual deve ser natural, lógico e de fácil compreensão. Isto significa que parâmetros como a potência pode ser demonstrada verticalmente durante a sua variação; a intensidade através do tamanho dos

⁷¹ ÖSTER, A. *Clinical Application of Computer-Based speech training for children with hearing impairment*. p.1

objetos, tornando-os grandes quando a intensidade está alta e diminuindo-os quando a intensidade está baixa; a entonação pode ser demonstrada através de curva; a duração pode ser mostrada horizontalmente; e a voz através da relação entre voz e mudança de cores;

- o programa deve possibilitar a comparação simultânea entre a locução da criança e a locução ideal;
- o software deve ser flexível, individual e deve dar uma avaliação dos resultados do treinamento da criança.

Segundo Pritsch⁷², o *feedback* visual pode ser dado de quatro formas:

- pelo o *sinal de tempo*. Exibem-se as características típicas das vogais (grande energia e sinais periódicos) e sons vocais (baixa energia e barulho) mantendo a informação (visual) muito restrita. As mudanças de fase, que acontecem facilmente quando o som é processado por sistemas eletrônicos, podem causar grandes mudanças nas formas dos sinais, representando algo que não simula a realidade, uma vez que estas mudanças de fase não se diferenciam facilmente pelos ouvidos;
- o *sinal de frequência*. Representado por espectrogramas, descrevem melhor a informação da fala e são muito utilizadas nesse tipo de software. A correlação entre as estruturas formadas e a audição é grande.
- a terceira representação é baseada no *poder de informação* do sinal de frequência. É similar ao espectrograma e pode ser associado a ele. Normalmente utilizam cores que definem marcas para as energias. É comum o uso do vermelho escuro para energias altas, e do amarelo para energias baixas. O espectro da fala é dividido em três diferentes frequências e representadas pelas cores vermelho, azul e verde. A mistura das cores na tela é baseada no poder de cada frequência. A altura de cada faixa depende da altura do sinal de voz;
- a quarta representação é a transformação direta da fala em uma imagem da *posição do movimento articulatório*, como pode ser visto no exemplo da Figura 17.

⁷² PRITCH, M. *Visual Communication Aid for Deaf People*.

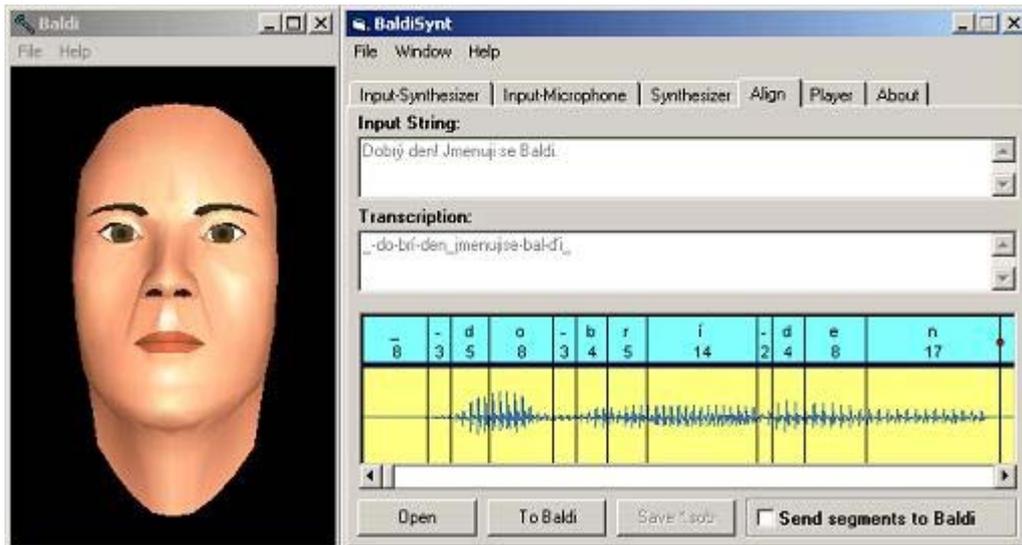


Figura 17 - Baldi. Software que exibe o movimento articulat3rio da fala no instante da pronuncia

Pritsch defende a quarta representa3o afirmando que os gr3ficos e espectrogramas s3o muito abstratos e n3o fornecem a express3o correta, nem a articula3o. Entretanto, a representa3o visual atrav3s de jogos e imagens l3dicas, acompanhadas das representa3es pict3ricas do movimento articulat3rio, pode auxiliar crian3as na idade de aquisi3o de linguagem brincando. Com isso 3 encontrada uma quinta forma de representa3o visual da fala que pode englobar todas as anteriores e possibilitar um car3ter l3dico 3 terapia.