

SOFTWARE PARA A AVALIAÇÃO DE CONHECIMENTO E AUXÍLIO AO APRENDIZADO DO IDIOMA INGLÊS ATRAVÉS DE INTERAÇÃO POR VOZ EM UM AMBIENTE VIRTUAL 3D

Caio Carvalho Moreira – caio.moreira@itec.ufpa.br

Elison Henrique Ribeiro Brito – outroelison@gmail.com

Manoel Ribeiro Filho – mrf@ufpa.br

Universidade Federal do Pará, Faculdade de Engenharia da Computação

Rua Augusto Corrêa, 01

66075-110 – Belém - Pará

Resumo: Atualmente, o inglês é a principal língua utilizada para a comunicação internacional, sendo considerada um dos bens simbólicos mais valorizados em nosso país. Assim, o conhecimento do idioma para aqueles que buscam boas qualificações no mercado de trabalho é imprescindível, pois esse conhecimento é requisito obrigatório na maioria das grandes empresas. Para obtê-lo, é necessário que o indivíduo pratique bastante e de forma eficiente, porém, devido a várias dificuldades como indisponibilidade de tempo, falta de investimento, timidez e difícil acesso às pessoas que conhecem o idioma, seu aprendizado pode se tornar uma tarefa árdua. Este trabalho, inicialmente, introduz os conceitos de duas tecnologias fundamentais para a compreensão do mesmo: realidade virtual e reconhecimento de voz. Em seguida, descreve o processo de criação e funcionamento de uma ferramenta educacional em forma de software, proposta para a avaliação de conhecimento e auxílio ao aprendizado do idioma inglês através de interação por voz com o computador e situado em um ambiente virtual 3D, sendo denominada *Virtual English School (VES)*. Este trabalho descreve, também, a metodologia utilizada para as avaliações e, por fim, ilustra através de gráficos comparativos e quantitativos os resultados obtidos, onde pode ser verificada sua atuação nos quesitos propostos: aprendizado e avaliação do inglês, podendo ser considerada relevante e confiável, respectivamente.

Palavras-chave: Escola virtual, Aprendizado, Avaliação, Realidade virtual, Reconhecimento de voz.

1 INTRODUÇÃO

O aprendizado de uma segunda língua, principalmente o inglês, vem se tornando de fundamental importância para conquistar espaço no mercado de trabalho e para se comunicar bem em qualquer parte do mundo. Segundo Rajagopalan (2005) e Moita Lopes (2005) citados por Tílio e Rocha (2009), na sociedade atual brasileira, o idioma inglês é imprescindível, tanto que hoje se tornou um dos bens simbólicos mais valorizados em nosso país, tendo um papel fundamental na transformação social. O Inglês é a principal língua usada para comunicação internacional e, por isso, é o idioma estudado por um maior número de brasileiros (MONTREZOR e SILVA, 2009).

A proposta deste trabalho consiste no desenvolvimento, funcionamento e avaliação de um *software* educacional, baseado em realidade virtual não imersiva, denominado *Virtual English School (VES)*. Esta ferramenta pode simular conversas cotidianas com o computador através de reconhecimento e síntese de voz e situado em um ambiente virtual tridimensional.

A justificativa da criação dessa ferramenta possui vários motivos como, disponibilizar a prática da conversação do idioma em horários livres, reduzir os custos do aprendizado, possibilitar uma maior interação com o idioma em função da diminuição da timidez, falta de investimento e dificuldade de acesso às professores e pessoas que conheçam o idioma

2 OBJETIVOS

O *software* VES tem como objetivo, conseguir através de sua utilização, uma melhora significativa na pronúncia e na compreensão do idioma inglês em estudantes de todas as idades e níveis, essa ferramenta pode simular uma conversa cotidiana, com o computador, em um ambiente virtual 3D, através do reconhecimento e síntese de voz, seguindo as etapas necessárias para que qualquer comunicação ocorra bem. Com o intuito de melhorar a pronúncia de cada palavra individualmente, o *software* o informa, sobre cada palavra, se a pronúncia precisa de um pouco mais de atenção, se está boa ou se está excelente, objetivando nas próximas vezes ter sua fala melhorada.

Também, a ferramenta propõe uma avaliação detalhada do nível de conhecimento do idioma do estudante. Cada um receberá uma média geral de seu desempenho, essa nota servirá de parâmetro para as futuras comparações, auto-avaliações e avaliações da ferramenta. Com essa quantificação do conhecimento, espera-se que os usuários se esforcem ao máximo para melhorar, cada vez mais, sua pronúncia.

3 REALIDADE VIRTUAL

A realidade virtual (RV) é um fenômeno da atualidade porque pode ser compreendida e aprendida muito facilmente (KIRNER & TORI, 2004). Seu leque de aplicações é enorme, podendo ser utilizada em diversas áreas e ciências como, na engenharia para a criação de protótipos, na arquitetura para a simulação de ambientes, na medicina para o treinamento de procedimentos e cirurgias, na educação com a interatividade do assunto estudado e o que o aluno pode ver na tela e em várias outras áreas e ciências, a depender da criatividade dos desenvolvedores.

4 RECONHECIMENTO DE VOZ

Para nós, humanos, reconhecer sons, identificar o que foi dito e quem disse geralmente é fácil, mas para um computador isto é um processo complexo (TAFNER, 1996) que envolve processamento de sinais e algoritmos de inteligência computacional. Este fato acontece porque nosso cérebro possui uma alta capacidade de desenvolvimento cognitivo, facilitando os processos de aprendizagem e aquisição de conhecimento, isto é, somos ótimos em perceber, se concentrar, armazenar informações, imaginar, raciocinar e etc. Tecnicamente, o cérebro é perfeito para execuções de tarefas não lineares. Já para os computadores, isto é um processo complexo, envolvendo processamento digital de sinais e inteligência computacional.

5 FERRAMENTAS UTILIZADAS

5.1 Ogre3D

Ogre3D é uma *engine* gráfica, ou seja, um motor gráfico onde sua principal função é facilitar o desenvolvimento de diversos aplicativos, como jogos, ferramentas de treinamento, simuladores ou animações que utilizam gráficos 3D acelerador por hardware. A classe da biblioteca abstrai os detalhes de se utilizar as bibliotecas gráficas de mais baixo nível como Direct3D e OpenGL, e também provê uma arquitetura com o paradigma da programação

orientada a objetos e outras classes intuitivas (BARATA, 2010). As principais vantagens desse motor gráfico, determinantes na escolha da mesma para o desenvolvimento do *software* foram, primeiramente, por ser a mais poderosa biblioteca de código aberto de renderização em tempo real da atualidade com isso evitaram-se custos na compra de licenças para sua utilização, além do mais, as familiaridades devido a trabalhos anteriores ajudaram a concluir que a *engine* é madura, estável, flexível, multi-plataforma (JUNKER, 2006).

5.2 3dsmax

3dsmax é um *software* de desenho 3D capaz de criar geometrias e animações virtuais, assim como definir as texturas e materiais utilizados nos mesmos, possibilitando uma visualização realística de um mundo virtual (MURDOCK, 2003). Este *software* é proprietário e está presente no mercado desde a década de 1990, voltado principalmente para a criação de filmes de animação e de objetos 3D para jogos (NASCIMENTO, 2010). Com esta ferramenta é possível criar todo um ambiente virtual e animá-lo sem a necessidade da *engine* gráfica, porém, a grande maioria das funções de interação está presente e só podem ser utilizadas através do motor gráfico.

5.3 SAPI

Speech Application Programming Interface (SAPI) é a *Application Programming Interface* (API) desenvolvida pela *Microsoft* que possibilita o reconhecimento e síntese de voz pelo computador. Existem dois tipos básicos de sistemas, que são o *Text-To-Speech* (TTS), ou seja, texto para fala, e *Speech Recognizer* (SR), isto é, reconhecedor de voz. O TTS sintetiza cadeias de texto e arquivos em áudio falado usando vozes sintéticas e o SR converte o áudio falado por humanos em cadeias de texto legível e em arquivos. Através das funções embutidas na API, é possível quantificar os resultados de cada palavra reconhecida, medindo um nível de confiança que pode ser baixo, normal ou alto, além de um número real no intervalo de $[-1,1]$ e também, existe a capacidade de personalizar a síntese de voz, escolhendo qual a voz, tonalidade, volume, velocidade e ênfase nas frases ditas.

5.4 AIML

A geração da resposta é baseada no modelo de linguagem *Artificial Intelligence Markup Language* (AIML), uma linguagem no formato *Extensible Markup Language* (XML) fácil de aprender e utilizar, tornando possível a modificação e criação de conversas rapidamente. Para que esta linguagem possa ser implementada, é preciso que haja um “motor” que faça a interface do *software* com a AIML, e neste caso é o Alicebot AIML, o mesmo utilizado no projeto *Artificial Linguistic Internet Computer Entity* (ALICE), um premiado bate-papo robô de inteligência artificial para linguagem natural livre.

A base de dados original do robô foi totalmente modificada para o interesse do desenvolvimento no sentido de desempenho, tamanho e simplicidade. Então, cinco diálogos compatíveis com os arquivos de gramática definidas na SAPI foram adaptados aos arquivos da base de dados do robô. Esses diálogos foram retirados e adaptados de um livro denominado “Fale Tudo em Inglês!” do autor brasileiro José Roberto A. Igreja, 2007, este tem um cunho didático para o aprendizado do idioma, onde o foco são as conversas situacionais e cotidianas, para que com isso haja um aprimoramento mais eficaz em situações comuns no nosso dia-a-dia.

6 DESENVOLVIMENTO DO SOFTWARE

6.1 Modelagem Tridimensional

Utilizando o 3dsmax, os cuidados tomados durante a modelagem são basicamente em relação ao número de faces, ou seja, a quantidade de polígonos e vértices criados para representar a geometria. Existe um custo computacional que aumenta proporcionalmente com o incremento faces, porém, é preciso levar em consideração a qualidade do modelo, que também aumenta com este número até o ponto onde não é mais diferenciável de um modelo com menor quantidade, elevando assim, o custo computacional desnecessariamente, logo, o modelador tem de ter sempre em mente o compromisso de quantidade de faces VS custo computacional VS aparência.

6.2 Configuração dos Diálogos

Gramáticas de Reconhecimento

Durante o desenvolvimento do *software* foi cogitado a possibilidade do uso de ditado livre, ampliando a abrangência da ferramenta a diversos assuntos, porém, os testes realizados levaram a conclusão que sua prática não levaria a uma boa experiência do usuário devido a enorme dificuldade em sua utilização. Assim, a melhor opção foi limitar a gramática de reconhecimento, o que aumenta imensamente as taxas de acerto e possibilita a avaliação de cada palavra reconhecida. Como dito, foram criados cinco diálogos sobre assuntos diferentes e três opções de fala para cada resposta do robô, a Figura 3 mostra essas opções.

Subject: Breaking the ice.
1. My family moved to Florida when I was just three.
2. My family moved to Florida when I was just a child.
3. I moved to Florida when I was a little boy.

Figura 3 – Exemplo das três escolhas para fala

6.3 Notas Obtidas

A avaliação do conhecimento de um segundo idioma é, com certeza, algo subjetivo, pois não existe uma única fórmula que, corretamente, nos dê um resultado satisfatório, surgiu então a necessidade inerente de quantificar os resultados para tal propósito. O *software* VES tem a capacidade fazer essa quantificação a cada palavra dita através de uma função que faz parte da API de reconhecimento de voz. O cálculo desses valores é feito pelo motor de reconhecimento de voz (SR), que retorna um valor de confiança, denominado *EngineConfidence*, como uma propriedade para cada evento gerado. Com isso, é possível avaliar o estudante atribuindo notas, por exemplo, de zero a cem, como mostra a Figura 4, além da possibilidade do estudante saber qual palavra foi pronunciada de forma excelente, boa ou má, como pode ser visto na Figura 5.

<p>Maior Nota 81.63125</p> <p>Ultima Frase 58.63258</p>	<p>Menor Nota 41.29042</p> <p>Media Geral 61.16396</p>
---	--

Figura 4 – Menu de notas

Falou Muito Bem: No. ; I ; was ; born ; South ; a ;
Falou Bem: in ; Carolina. ; but ; I ; live ; here ; since ; I ; was ; kid. ;
Precisa Melhorar:

Figura 5 – Menu de auxílio à correção da pronúncia

7 FUNCIONAMENTO

Inicialmente o estudante encontra-se do lado de fora do prédio da escola virtual, onde ele pode se locomover livremente pelo ambiente virtual. Durante esta etapa o estudante é sugerido a ir até a sala, dentro da escola, onde a professora se encontra, navegando com as setas do teclado e com o mouse. Também é sugerido verificar as funcionalidades da ferramenta através do comando de voz “*Help*”, como mostra a Figura 6.



Figura 6 – Ambiente Inicial do VES

7.1 Treinamento

A etapa seguinte inicia-se com as opções dos diálogos disponíveis na parte inferior da tela, onde o usuário pode escolher entre cinco opções, como é visto na Figura 7. O estudante então pode conversar com o computador, respondendo cada afirmação ou pergunta com uma das três opções disponíveis naquele momento. Ao terminar aquele assunto, ele é novamente direcionado à janela de opções dos diálogos, onde pode escolher outro assunto ou até o assunto abordado novamente, se desejar repeti-lo, este processo pode se repetir indefinidamente, deixando-lo livre para utilizar a ferramenta como bem entender.

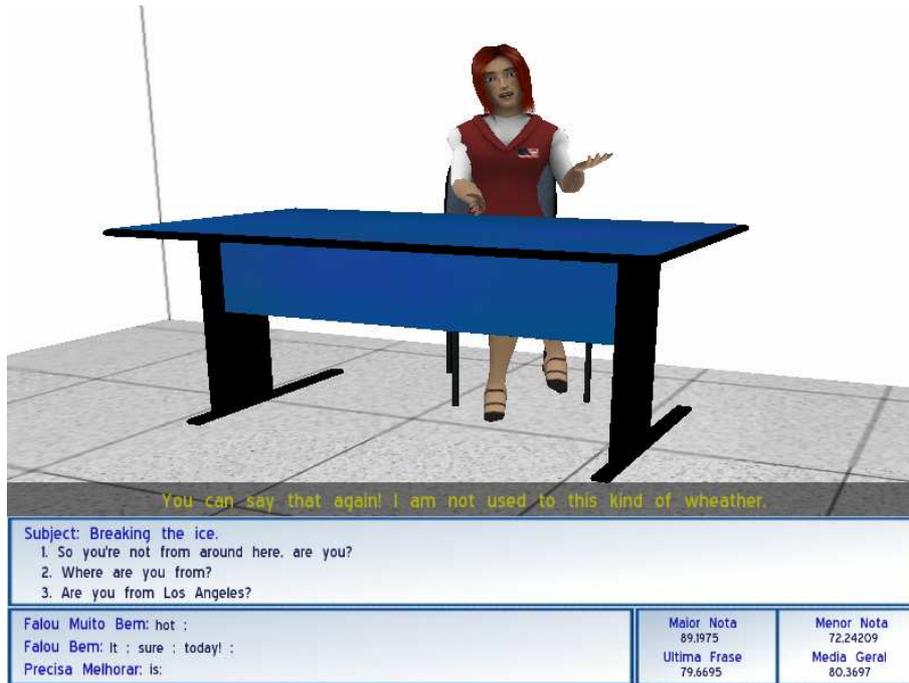


Figura 7 – Conversação

7.2 Resultados Finais

Ao executar o comando de sair do programa, *Quit Program*, disponível no menu de funcionalidades *Help*, é apresentado ao estudante uma janela onde são mostrados os dados de seu desempenho. Estes dados são: Maior nota, Menor Nota e Média Geral. O *software* também o classifica, de acordo com a média obtida, como: Fluente, Avançado, Intermediário, Iniciante ou Sem Contato com o Idioma, , como mostra a Figura 8. O estudante finaliza o programa pressionando a tecla escape ou “*Esc*” do teclado.

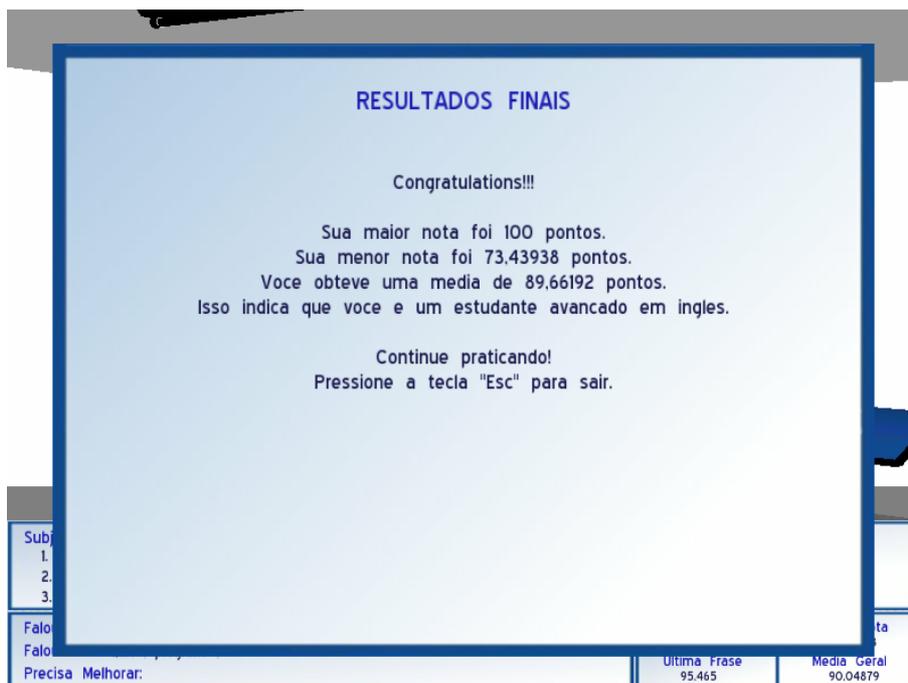


Figura 8 – Resultados finais

8 AVALIAÇÃO

8.1 Metodologia

A metodologia utilizada para a avaliação consistiu em duas partes, a primeira sendo uma auto-avaliação em relação ao conhecimento do inglês, feita antes da utilização do *software*, e a segunda sendo uma pesquisa de satisfação da ferramenta e a obtenção das notas obtidas pelos estudantes, feita após a utilização da ferramenta. Foram avaliados vinte estudantes, graduandos, de dezoito à vinte seis anos de idade, com níveis de conhecimento variado em inglês. Os questionários possuem três objetivos principais: avaliação e relevância da ferramenta, a averiguação o grau de precisão na avaliação do estudante feita pelo *software*.

8.2 Resultados

Comparativos

Como a média geral, na verdade, avalia somente a pronúncia do usuário e, considerando que alguns estudantes têm mais facilidade ou dificuldade em outros quesitos, estas informações foram relacionadas às informações do quesito sobre habilidade em pronúncia, respondido na auto-avaliação, como mostra a Figura 9.

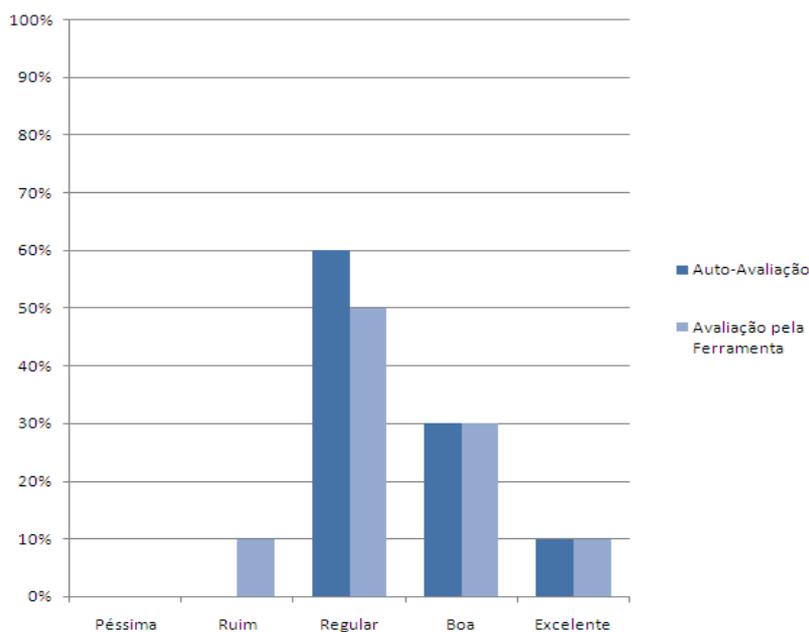


Figura 9 – Nível atual de pronúncia VS Média geral

Pode-se perceber que a ferramenta avalia com confiança o inglês do estudante, pois as médias dos resultados de pronúncia boas e excelentes obtiveram o mesmo percentual e, embora os resultados regulares e ruins não alcancem o mesmo feito, a divergência entre eles foi de apenas 10%.

Quantitativos

Um gráfico importante que mostra a satisfação dos usuários em relação à relevância que a ferramenta teria melhora do inglês, isto é, se o estudante a utilizasse durante certo tempo, na opinião dele, o quão benéfico o seria, em relação à pronúncia, é mostrado na Figura 10.

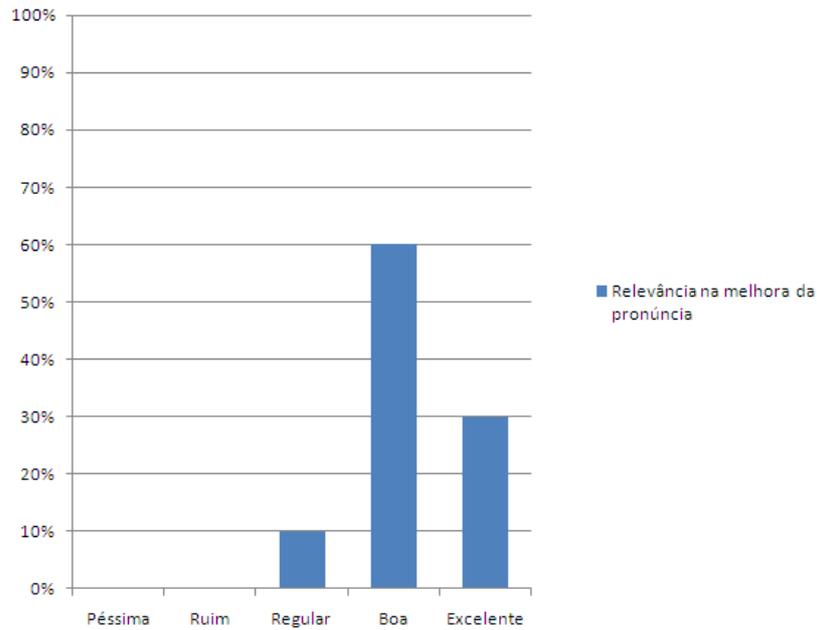


Figura 10 – Relevância na melhora da pronúncia

Pode-se perceber uma aprovação de 90% dos estudantes, mostrando que os estudantes acreditam em uma possível melhora em sua pronúncia com a utilização da ferramenta, assim, este é um resultado expressivo que colabora para o objetivo deste trabalho.

Por último, uma pesquisa de satisfação que leva em consideração todos os seus aspectos, ou seja, uma pesquisa de satisfação geral sobre o *software*, sua utilização e relevância, como mostra a Figura 11.

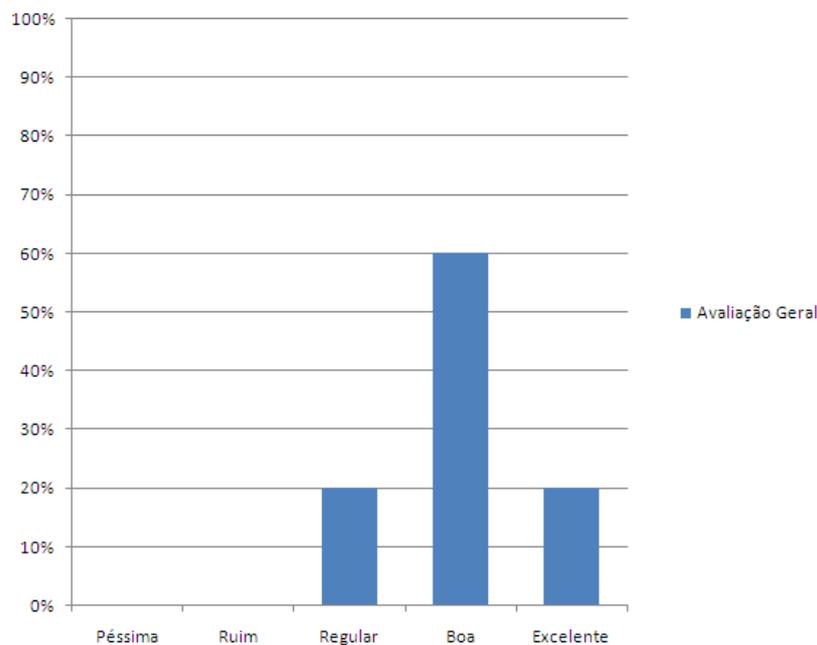


Figura 11 – Satisfação geral

É possível observar que o percentual de aprovação é de 80%, o que provavelmente satisfaz os estudantes em relação a sua experiência de utilização e seu aprimoramento do idioma inglês.

9 CONCLUSÃO

Com esse índice de aprovação, os objetivos foram atingidos, possibilitando aos estudantes uma melhoria na compreensão e, principalmente, na pronúncia com a utilização da ferramenta durante um período adequado.

Assim, pode-se considerar que o *software* tem relevância e sentido de existir. Sua utilização trará benefícios aos estudantes, que poderão aperfeiçoar seu domínio sobre o idioma inglês com um maior contato com a ferramenta.

Porém, é importante lembrar que esta *software* não substitui o professor em hipótese alguma, visto que é uma ferramenta para auxílio ao aprendizado e não para a formação integral de conhecimento.

10 REFERÊNCIAS / CITAÇÕES

BARATA, Pebertli Nils Alho. UNIVERSIDADE FEDERAL DO PARÁ, Instituto Tecnológico. Projeto e implementação de um Sistema de Autoria para animações, simulações e treinamentos em Realidade Virtual, 2010. 22 p, il. Dissertação (Mestrado).

IGREJA, José Roberto A. Fale Tudo em Inglês!, 4. ed. São Paulo: Disal, 2007. 6 p, il.

JUNKER, Gregory. Pro OGRE 3D Programming. 1. ed. New York: Apress, 2006. 19 p, il.

KIRNER, Claudio; TORI, Romero. Realidade Virtual: Conceitos e Tendências. VII Symposium on Virtual Reality, Centro Universitário SENAC-SP, São Paulo, SP, 2004.

MONTREZOR, Bethania Márcia; SILVA, Alexandre Batista da. A dificuldade no aprendizado da Língua Inglesa. **Cadernos UniFOA**. Volta Redonda, ano IV, n. 10, p1-4, 2009.

MURDOCK, K. L. La Biblia de 3ds max 5. Madrid: Anaya Multimedia, 2003. 9 p, il.

NASCIMENTO, Messias José Amador do. UNIVERSIDADE FEDERAL DO PARÁ, Instituto Tecnológico. Modelagem de Objetos Tridimensional para um Ambiente Interativo de Instruções Técnicas Virtuais, 2010. Dissertação (Mestrado).

SILVA, Romano José Magacho da; RAPOSO, Alberto Barbosa; GATTAS, Marcelo. Grafo de Cena e Realidade Virtual. **Monografias em Ciência da Computação**, n 11/04, p3-8, 2004.

TAFNER, M. A. UNIVERSIDADE FEDERAL DE SANTA CATARINA, Centro Tecnológico. Reconhecimento de Palavras Faladas Isoladas usando Redes Neurais Artificiais. 1996, 15 p, il. Dissertação (Mestrado)

TÍLIO, R.; ROCHA, C. H. “As dimensões da linguagem em livros didáticos de inglês para o Ensino Fundamental I”. **Trabalhos em Linguística Aplicada**, v. 48, Unicamp, Campinas, SP, 2009. ISSN 0103-1813.

TORI, Romero; KIRNER, Claudio; SISCOOTTO, Roberto Augusto. Fundamentos e Tecnologia de Realidade Virtual e Aumentada. VIII Symposium on Virtual Reality, Belém, PA, 2006.

TRIGUEIRO, Osvaldo. O estudo científico da comunicação: avanços teóricos e metodológicos ensejados pela escola latino-americana. **PCLA Revista Científica Digital**, v. 2, n. 2, 2001.

SOFTWARE FOR EVALUATION OF KNOWLEDGE AND AID TO LEARNING ENGLISH USING VOICE INTERACTION IN A VIRTUAL 3D ENVIROMENT

***Abstract:** Currently, English is the main language used for international communication and is considered one of the most valued symbolic goods in our country. Thus, knowledge of the language for those who seek well-educated labor market is essential, since this knowledge is a mandatory requirement in most large companies. To get it, it is necessary that the individual practices a lot and efficiently, however due to various difficulties such as lack of time, lack of investment, shy and difficult access to people who know the language, their learning can become a arduous task. This paper initially introduces the concepts of two key technologies for the understanding of it: virtual reality and speech recognition. It then describes the process of creating and operating an educational tool in the form of software proposed for the evaluation of knowledge and assistance to English language learning through interaction with the computer voice and set in a 3D virtual environment, being called Virtual English School (VES). This paper also describes the methodology used for evaluation and, finally, illustrates through quantitative graphs and comparative graphs the results obtained, which can be verified its performance in the categories proposed: learning and assessment of English and may be considered relevant and reliable, respectively.*

Key-words: Virtual school, Learning, Evaluation, Virtual reality, Voice recognition.