

# Abordagem empírica para avaliação de uma aplicação de exercício linguístico utilizando interface de voz

Douglas Fabris Aguiar<sup>1</sup>, Marcos Arrais e Silva<sup>1</sup>

<sup>1</sup>Instituto de Educação Continuada - Pontifícia Universidade Católica de Minas Gerais (PUC Minas) Belo Horizonte - MG - Brasil,

{devfabris, marcos.arrais}@gmail.com

**Abstract.** *Linguistic exercise applications have become widely known and present in people's daily lives due to their ease of access. This presence and relevance brings critical views regarding the pedagogical alignment of the activities proposed by the applications and opens space for investigation and hypothesis of the use of voice interfaces in this scenario. With this in mind, this work aims to create a proposal for the application of a linguistic exercise and a heuristic evaluation based on the Task-based language learning method to identify whether there are benefits in using this type of interface for foreign language practice.*

**Resumo.** *Aplicações de exercício linguístico têm se tornado amplamente conhecidas e presentes no dia a dia das pessoas devido a sua facilidade de acesso. Essa presença e relevância traz olhares críticos quanto ao alinhamento pedagógico das atividades propostas pelas aplicações e abre espaço para a investigação e hipótese do uso de interfaces de voz nesse cenário. Tendo isso em mente, o presente trabalho objetiva a criação de uma proposta de aplicação de exercício linguístico e uma avaliação heurística baseada no método Task-based language learning para identificar se existem benefícios no uso do tipo de interface para prática de idioma estrangeiro.*

## 1. Introdução

A tecnologia tem mudado vários aspectos da vida humana e com o aprendizado de idiomas isso não é diferente. Essa inaugurou uma nova era de aprendizagem, tornando o ensino envolvente, flexível e interessante, aumentando a produtividade e eficiência [Halverson and Smith, 2009].

No passado, a interação com computadores dava-se por meio de uma interface utilizando comandos pré-estabelecidos acessível somente a pessoas com conhecimento técnico. Atualmente, as interfaces evoluíram, possibilitando o seu uso por um grande número de pessoas com pouco conhecimento técnico [Grudin J, 2012].

Interfaces que vêm se destacando nas últimas décadas são as de voz e a síntese de fala. Por meio delas, as conexões tornaram-se acessíveis para inúmeras pessoas, incluindo deficientes visuais. Provando que a troca de informações feita de modo verbal em situações de interação homem-máquina é cerca de duas vezes mais eficiente do que qualquer outra forma de comunicação [Chapanis A, 1988].

Aplicações que auxiliam no aprendizado de novos idiomas têm sido cada vez mais populares, de fácil acesso e apresentam relevância no auxílio do desenvolvimento de estudos. Porém, existem preocupações de que grande parte destas não estejam alinhadas com a teoria da aquisição de um segundo idioma [Matthews and Burke, 2019].

Tendo em vista esse desalinhamento pedagógico das atividades presentes em aplicativos para aprendizado de idiomas e evidências anteriores em relação à como adquirimos uma segunda língua e ao uso, cada vez mais popular, de interfaces de voz [Aguiar e Arrais, 2020], torna-se oportuno aprofundar o conhecimento em relação à essas falhas e identificar meios de avaliar interfaces de voz para verificar se existem vantagens no uso desse tipo de interface em aplicações de exercício linguístico.

Diante dessa problemática, este trabalho propõe a produção de um protótipo de um aplicativo de exercício linguístico focado na língua inglesa, utilizando interfaces de voz e uma inspeção heurística baseada em princípios do método *Task-based language learning* (TBLT).

## **2. Referencial Teórico**

Atualmente, existe um número grande e crescente de aplicativos de aprendizagem de idiomas que podem facilmente serem baixados, instalados e usados. Levando em consideração essa disponibilidade, o potencial de aproveitar e entregar funções de tecnologia móvel para fins de aprendizagem de idiomas por meio desses aplicativos é uma perspectiva empolgante [Matthews and Burke 2019].

Apesar desse potencial aparente, a realidade do uso desses aplicativos no ensino de um segundo idioma tem sido frustrante em alguns aspectos [Reinders and Pegrum, 2015], visto que ainda existem alguns déficits, principalmente do ponto de vista pedagógico do ensino de idiomas estrangeiros [Burston, 2014].

### **2.1. Interfaces de voz e sistemas de diálogo falado**

As interfaces por voz são alternativas interessantes para sistemas cuja interação é genericamente feita através de entrada de texto, cursor de mouse e dispositivos touch screen. O uso da fala na interface de um sistema computacional tem inúmeras vantagens, como diminuição do tempo de formação e adaptação do utilizador, aumento de produtividade e ampliação do universo de utilização a pessoas menos especializadas ou com necessidades especiais.

Pode-se afirmar que os Sistemas de Diálogo Falado (SDF) são úteis para os sistemas computacionais atuais e do futuro, esses estão mudando o paradigma de comunicação entre pessoas e computadores, aumentando os modos de interação de uma forma excepcional. Esses sistemas consistem em um software que utiliza a representação lógica de relações, sequências e textos de linguagem humana para simular e reconhecer a fala, baseada numa interação por áudio.

Ao se depararem com dispositivos capazes de interagir pela fala, as pessoas esperam ter uma conversa natural e desejam que o computador os entenda, o que, por vezes, torna-se um impedimento para que este consiga desempenhar uma tarefa. Como tal, existem fatores que devem ser levados em consideração quando se pretende criar uma interface conversacional, como os aspectos emocionais e cognitivos do utilizador, a complexidade, eficiência e eficácia da tecnologia e o papel das pessoas no sistema, bem como a sua memória.

Para [Young et al., 2013] o entendimento da linguagem falada é responsável por separar o texto em segmentos funcionais de forma a definir os atos dialogais e a semântica de cada segmento. O gestor de diálogo é o responsável por escolher uma resposta apropriada às entradas de diálogos e por manter o diálogo e a tomada de decisão sempre que necessário.

Quando o utilizador fala, o sinal de áudio é convertido pelo ASR (*Automatic Speech Recognizer*) que posteriormente, se necessário, é tratado para extração do seu significado relevante para a aplicação pela SLU (*Spoken Language Understanding*), através de gramáticas ou métodos estatísticos [Wang, Ye-Yi, 2005]. A arquitetura cliente-servidor é a mais comum em um sistema conversacional. Nesta arquitetura de componentes temos um cliente simples (interface de voz) e um servidor responsável pelo gerenciamento do diálogo e comunicação da aplicação com o cliente.

Aspectos como a arquitetura das interfaces ou os modelos de fluxo do diálogo são essenciais ao desenvolvimento de uma boa interface de voz, mas precisa existir a aceitação dos usuários. É fundamental desenhar a interface de voz com base nas melhores práticas de usabilidade no domínio da voz e, para tal, é necessário conhecer essas práticas e regras [Weinschenk, 2000].

## **2.2. Avaliação de interfaces**

Avaliações são necessárias para responder as dúvidas que surgem durante o processo de design e desenvolvimento de uma solução ou um produto. Em muitos pontos do processo de design, as pessoas envolvidas precisam verificar se suas ideias são o que os usuários necessitam ou desejam. Desse modo, a avaliação direciona e mescla-se ao design, auxiliando na criação de um produto útil e usável. Resumidamente, podemos dizer que a avaliação tem três grandes objetivos: avaliar a funcionalidade do sistema, avaliar o efeito da interface junto ao usuário e identificar problemas específicos do sistema [Da Rocha e Baranauskas, 2003].

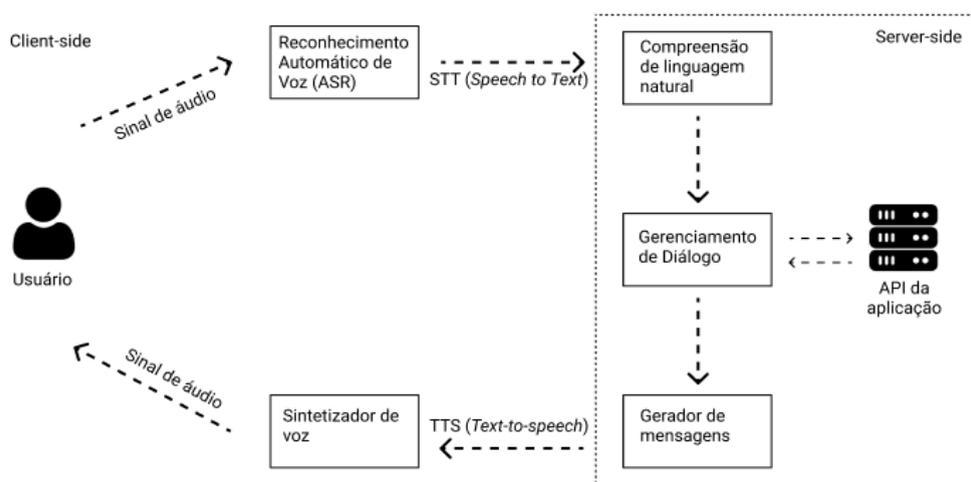
Contudo, apesar dos aspectos de usabilidade serem menos considerados, sabe-se que é um fator tão importante como a qualidade técnica dos componentes destes sistemas. Segundo [Dybkjaer and Bernsen, 2000], para que seja possível avaliar a qualidade de um sistema de reconhecimento de voz é necessário responder questões como: o que está a ser avaliado, qual é o tipo de avaliação, sintomas a procurar, importância da avaliação, dificuldade da avaliação, custo da avaliação e ferramentas de suporte à avaliação. De acordo com o autor, a resposta a este tipo de perguntas permite que a avaliação da interface seja eficaz e eficiente. É importante salientar que este tipo de avaliação pode variar com as necessidades dos sistemas ou até mesmo do utilizador ou grupo de utilizadores do mesmo.

É importante considerar e planejar os requisitos de usabilidade do projeto para que as necessidades e capacidades dos utilizadores sejam levadas em consideração. Para [Möller et al., 2008] as falhas nos SDF normalmente ocorrem quando há divergências entre o ponto de vista do fornecedor e do utilizador.

### 3. Metodologia

#### 3.1. O protótipo da interface conversacional

Para identificar e avaliar se de fato existem benefícios na utilização de interfaces de voz foi proposto um protótipo de um aplicativo, utilizando interface conversacional e simulando atividades de exercício linguístico, em que o usuário pode interagir e, baseado nas respostas, o sistema faz o roteamento das atividades. A figura 1 mostra a arquitetura do protótipo e as suas inter-relações:



**Figura 1. Arquitetura do protótipo da interface conversacional.**

O protótipo funciona da seguinte forma: o usuário inicia a aplicação através de um dispositivo ativado por voz, a partir dessa ativação, cada entrada do usuário é analisada pela aplicação que verifica o nível de assertividade (Compreensão de linguagem natural) daquela entrada. Caso a aplicação compreenda a entrada, é feito um roteamento para redirecionar o usuário (Gerenciamento de Diálogo). Caso o resultado do processamento da entrada não seja satisfatório, o sistema faz tentativas de extrair do usuário uma nova entrada para tentar direcioná-lo para outras partes da aplicação.

A principal funcionalidade da aplicação é o reconhecimento da fala do usuário utilizando ASR, que, por sua vez, extrai a necessidade do usuário de tentar ser entendido, fazendo com que o aluno de idioma estrangeiro tenha que se esforçar ao se expressar com a interface de voz. O protótipo foi criado utilizando a *Dialog Engine* da plataforma VoiceFlow<sup>1</sup> que possui uma interface visual para montagem do protótipo e os fluxos de interação, em que é possível integrar e disponibilizar para testes em assistentes consolidados, como Google Assistant e Amazon Alexa.

#### 3.2. Método de avaliação

Uma descoberta abrangente foi a de que muitas das atividades de instrução oferecidas pelos aplicativos de aprendizado de idiomas pareciam ser fundamentalmente do modelo de transmissão de instrução do tipo comportamentalista, conforme presumido por [Burston, 2014]. Ficou evidente a partir das descrições dos professores das aplicações

<sup>1</sup> Voiceflow: <https://www.voiceflow.com/>

que a instrução enfatizou predominantemente a análise metalinguística e o foco descontextualizado nas formas, e isso vai contra a premissa básica do *Task-based language learning* (TBLT) proposta por [Ellis, 2003].

No contexto do TBLT, [Ellis 2003] define-se uma tarefa como uma atividade de ensino de línguas que coloca um foco principal no significado e requer que o aluno compreenda e transmita mensagens comunicativas na língua-alvo. Ellis afirma ainda que uma tarefa requer que o aluno aplique seus próprios recursos linguísticos e não linguísticos para negociar uma lacuna de informação a fim de alcançar um resultado comunicativo. Essa percepção e o conhecimento do método, encoraja o afastamento do design instrucional que se concentra predominantemente em estruturas linguísticas e, em vez disso, enfatiza tarefas que exigem que os alunos usem a linguagem para negociar o significado e alcançar objetivos comunicativos [Doughty and Long 2003].

Para realizar a avaliação do protótipo desenvolvido, foi utilizado a avaliação heurística, um método de inspeção de usabilidade. Essas heurísticas foram baseadas na interpretação da estrutura de [Doughty and Long 2003] para TBLT, uma vez que aborda explicitamente a realização de TBLT mediada por computador e pelo fato de ser uma estrutura que fornece dez heurísticas metodológicas do modelo de aprendizagem e ensino de línguas com base em tarefas. Essas heurísticas são apresentados na tabela 1.

**Tabela 1. Heurísticas do modelo de aprendizagem e ensino de línguas com base em tarefas.**

<b>Foco no uso de tarefas como análise</b>	O foco do aluno não está na análise da linguagem, mas na formulação de frases e significados socialmente contextualizados. O conteúdo das atividades deve ter evidência na conclusão de tarefas.
<b>Aprendizado ativo</b>	Ajudar os alunos a atenderem às suas necessidades comunicativas do mundo real, por meio do incentivo a atividades práticas relativamente autênticas de resolução de problemas.
<b>Formulação da linguagem</b>	Evitar a necessidade de formulação excessiva da linguagem e focar no significado dos textos e frases autênticas.
<b>Fornecimento de conteúdo</b>	Garantir que os alunos completem tarefas por meio de envolvimento com textos que estão alinhados com as necessidades de aprendizagem de línguas dos alunos.
<b>Aprendizagem indutiva</b>	Ao invés de estudar o idioma por meio da análise explícita de amostras de linguagem, oferecer oportunidades para que a obtenção do conhecimento seja por meio da indução.
<b>Foco no significado</b>	Facilitar para que a atenção dos alunos seja direcionada para itens linguísticos que permitam o engajamento em tarefas interativas com foco no significado.
<b>Fornecimento de feedback</b>	Fornecer feedback negativo explícito ou implícito em resposta à produção de uma forma linguística errônea.
<b>Desenvolvimento individual</b>	Os objetivos das tarefas a que os alunos são expostos precisam ser formulados por meio de uma análise contínua, evolutiva e individualizada da capacidade dos alunos.
<b>Aprendizado colaborativo</b>	Facilitar o discurso entre os alunos à medida que eles evoluem em tarefas, de forma colaborativa e cooperativa.
<b>Instrução de individualização</b>	Garantir que a aprendizagem seja adaptada às diferenças individuais, e isso pode ser alcançado por meio de análises de necessidades.

Para realizar a avaliação, foram convidados outros dois profissionais com experiência prévia em avaliação heurística, sendo um deles especialista em desenvolvimento de aplicações para interfaces conversacionais. Para esta tarefa, os avaliadores foram convidados a interagir com o protótipo que possuía uma atividade de exercício linguístico em que simulava um diálogo em uma pizzeria. A figura 2 mostra um exemplo de interação com o protótipo.

S: Imagine you're at an Italian restaurant and want to eat a pizza.  
S: Hello, welcome to Mr. Pizza, how can I help you?  
U: I want to eat a pizza.  
S: Hmm, you want a pizza, right? What flavor?  
U: Pepperoni.  
S: Ok, Pepperoni! I'm going to prepare for you...  
S: Do you want to drink something?  
U: Yes.  
S: What do you want to drink?  
U: Orange Juice.  
S: Ok, juice then.  
S: Congratulations, you finished your activity!

**Figura 2. Um diálogo com o protótipo. S é o sistema, U é o usuário.**

Assim, a avaliação do protótipo foi composta por três etapas:

1. Elaborar uma ficha de avaliação que deve ser preenchida pelos avaliadores. O conteúdo desta ficha contemplou os princípios presentes na tabela 1 e os avaliadores tinham como opção de classificação: “Concordo”, “Não concordo”, “Não aplicável” relacionado a cada princípio;
2. Cada especialista avalia o protótipo verificando se foram cumpridos os princípios da abordagem, reportando as falhas e o nível delas, em relação ao compromisso do princípio, encontradas na aplicação;
3. Compilação dos resultados e criação de um resumo da avaliação com base nos dados coletados pelos especialistas envolvidos para apresentação e discussão.

### **3.3. Resultados e discussão**

As principais observações dos avaliadores com base na lista de heurísticas da tabela 1 são apresentadas abaixo:

1. Três dos quatro especialistas concordam que a proposta apresenta foco no uso de tarefas como análise pelo fato de tarefas serem a forma de apresentação dos exercícios linguísticos, porém deveria ser analisado junto aos usuários a recepção desse tipo de atividade.
2. Todos concordam que o protótipo promove o aprendizado ativo, por trazer situações cotidianas e rotineiras, além de ser o real foco da ideia da aplicação.
3. Dois especialistas concordam que o sistema evita o uso de formulação excessiva da linguagem. O protótipo é auto instrutivo, e tem a preocupação de fazer o usuário conseguir atingir seu objetivo em uma determinada situação do mundo real.
4. Todos discordam que o protótipo garante tarefas com conteúdo alinhado com as necessidades dos alunos, mas reconhecem o potencial e adicionam a necessidade

e importância de professores de idiomas envolvidos no processo de ideação dessas atividades.

5. Dois dos quatro especialistas concordam que, ao fazer que o aluno precise negociar seu conhecimento, para obter resultados do sistema, um processo indutivo acontece. Os outros dois especialistas acreditam que essa negociação não é o suficiente para a troca e aprendizado.
6. Todos os especialistas concordam que o sistema tem grande potencial em facilitar que a atenção dos alunos seja direcionada à itens que permitam atividades focadas no significado, mas, para que isso aconteça, as situações vivenciadas devem ser adaptadas para os interesses particulares dos alunos.
7. Os especialistas acreditam que, o princípio de feedback negativo, não se aplica ao sistema. Relataram que o feedback apresentado é apenas parcial e que esse, é um ponto que novamente existe a necessidade de validação com usuário. Os alunos precisam de feedbacks mais precisos, direcionados e não somente feedbacks que contornam erros não esperados.
8. Três dos especialistas discordam que o protótipo é adaptável às necessidades dos alunos. E que designers e profissionais envolvidos devem levar em consideração essa necessidade no planejamento de atividades.
9. Em sua totalidade os avaliadores acreditam que o princípio de promoção colaborativa não se aplica, mas que tem um grande potencial para que os estudantes troquem informações sobre seus aprendizados e experiências.
10. Para os avaliadores, para garantir o princípio relacionado às necessidades individuais seria necessário ter um repertório maior de testes e por isso responderam que o princípio não se aplica.

#### **4. Conclusão**

A avaliação heurística baseada nos princípios do TBLT trouxe uma abordagem diferente do convencional para avaliação de interfaces. A atenção com aspectos didáticos das aplicações possibilitou identificar pontos que não seriam levantados se trouxéssemos uma avaliação heurística mais convencional.

Foi levantado pelos especialistas problemas para executar o protótipo e outros relacionados ao funcionamento e reconhecimento de fala. Devido a isso e baseado nos resultados, fica evidente a necessidade de testes de usabilidade monitorados com usuários para validar pontos importantes e validação de alguns aspectos mais relacionados à aplicação e as atividades.

É oportuno dizer que, como trabalhos futuros, algumas atividades podem ser propostas, entre elas uma prototipação em mais alto nível, com participação de profissionais e professores de línguas estrangeiras para elaboração das atividades e um sistema para checar a pronúncia dos estudantes dando *insights* para o aluno quanto ao seus *inputs* em suas tarefas. Assim como foi constatado anteriormente por [Aguiar e Arrais, 2020] o avanço desse tipo de interface é bastante animador, mas o desafio para fazer com que as aplicações satisfaçam os desejos de usuários, principalmente aprendizes, é latente. Uma avaliação de uma aplicação mais avançada com o usuários se faz necessária, como foi relatado em vários aspectos pelos avaliadores deste trabalho,

confirmando que a interação humano-computador realiza-se com a presença dos usuários em várias partes do projeto.

## **Referências**

- Halverson, R. Smith, A. (2009) "How new technologies have (and have not) changed teaching and learning in schools". *Journal of Computing in Teacher Education*.
- Grudin, J. (2012) "A moving target - the evolution of human-computer interaction. In *Human-Computer Interaction Handbook (3rd Edition)*".
- Chapanis, A. (1988) "Interactive human communication. In *Computer-supported Cooperative Work: A Book of Readings*. Morgan Kaufmann Publishers Inc".
- Matthews, J. Burke, R. (2019) "Listening to Teachers' Voices: Evaluating Language Learning Apps with a Task-based Language Teaching Framework".
- Aguiar, D. F. Arrais, M. (2020) "Uma investigação sobre o uso de interfaces de voz como ferramenta para prática de idioma estrangeiro".
- Inácio Jr, V. R. (2007) "Um framework para desenvolvimento de interfaces multimodais em aplicações de computação ubíqua".
- Reinders, H. Pegrum, M. (2015) "Supporting language learning on the move: An evaluative framework for mobile language learning resources".
- Burston, J. (2014) "The reality of MALL: Still on the fringes". *Calico Journal*.
- Young, S. Gašić, M. Keizer, S. Mairesse, F. Schatzmann, J. Thomson, B. Yu, K. (2013) "POMDP-based Statistical Spoken Dialogue Systems: a Review".
- Wang, Y., Deng, L., Acero, A. (2005) "Spoken Language Understanding" *An Introduction to the Statistical Framework*".
- Weinschenk, S. Barker D.T. (2000) "Designing Effective Speech Interfaces".
- Da Rocha, H. V. Baranauskas, M. C. C. (2003) "Design e Avaliação de Interfaces Humano-Computador".
- Dybkjaer, L. Bernsen, N. (2000) "Usability issues in spoken language dialogue systems".
- Möller, S. Engelbrecht, K.-P. Schleicher, R. (2008) "Predicting the quality and usability of spoken dialogue services".
- Ellis, R. (2003) "Task-based language learning and teaching". Oxford, United Kingdom: Oxford University Press.
- Doughty, C. J. Long, M. H. (2003) "Optimal psycholinguistic environments for distance foreign language learning".