

Tecnologia de Transcrição da Fala como Ferramenta para Inclusão Digital na Educação

João Paulo Martins Barcelos Programa De Pós-Graduação Em Educação Profissional E Tecnológica Vitória, Brasil joaopmb@ifes.edu.br	Danielli Veiga Carneiro Sondermann Programa De Pós-Graduação Em Educação Profissional E Tecnológica Vitória, Brasil danielli@ifes.edu.br	Camila Petarli da Fonseca Programa De Pós-Graduação Em Educação Profissional E Tecnológica Vitória, Brasil camilapetarli@gmail.com	Evanilton Neri de Oliveira Programa De Pós-Graduação Em Currículo, Linguagens e Inovações (UFBA) Salvador, Brasil neri@ifes.edu.br
---	---	---	---

RESUMO

O uso de legendas como recurso pedagógico permite aos alunos com deficiência auditiva, aos surdos usuários da Língua Portuguesa e aos professores, dispensarem o uso da Língua Brasileira de Sinais e também a visualização da face do professor para a leitura labial. Este artigo teve como objetivo descrever as ferramentas de legendas simultâneas disponíveis por meio de *software* livre e/ou ambiente web e que podem ser utilizados no ambiente escolar, visando melhorar a aprendizagem de alunos com deficiência auditiva e surdos. Metodologicamente, este artigo pautou-se numa revisão bibliográfica exploratória, sistemática e qualitativa, cujo foco de pesquisa se dedicou a investigar a ocorrência de *softwares* e *sites* capazes de promover legendas simultâneas durante uma conversação. Os *softwares* identificados após a pesquisa bibliográfica foram o *Live Caption*, *Live Relay* e o *Web Captioner*; todos pertencentes ao projeto Euphonia, desenvolvidos pela empresa multinacional americana de serviços *online* e *softwares* Google.

Palavras-chave: Conversação; Aprendizagem; Alunos com deficiência auditiva e surdos; Tecnologia.

Palavras-chave de classificação do ACM: Computação centrada no ser humano; Acessibilidade; Tecnologias de acessibilidade.

INTRODUÇÃO

Durante a década de 1990 o entendimento sobre a inclusão digital que prevalecia no Brasil estava associado às questões socioeconômicas que os brasileiros enfrentaram no período. Nesse contexto, o baixo poder de compra dos indivíduos era um fator limitante para que fossem incluídos digitalmente já que a falta de recursos impedia a aquisição de produtos eletrônicos. Foi reduzido nessa época a um enfoque inclusivo por meio do uso da tecnologia visando a ampliação da diversidade cultural.

A inclusão digital associa-se não somente à aquisição de dispositivos como computadores, *tablets*, *smartphones*, entre outros. Embora os aparelhos sejam um mecanismo importante pelo qual o indivíduo se mantém conectado a diferentes interfaces, é por meio do acesso à rede e pelo

domínio das ferramentas tecnológicas que as pessoas deixam de ser excluídas digitalmente.

O uso de tecnologias digitais nas escolas permite interações socioculturais por meio da educação, o que pode viabilizar o acesso igualitário a dispositivos eletrônicos entre todos os alunos, independentemente de sua condição social e favorecer a inclusão escolar de alunos com deficiência, uma vez que a ferramenta digital é utilizada como recurso pedagógico. É essencial para a Educação Profissional e Tecnológica (EPT), porque visa uma maior integração dessas pessoas, e torna o processo mais igualitário, isonômico, receptivo, diminuindo assim a desigualdade e promovendo maior interação.

Os alunos com deficiência auditiva e surdos, que já estão alfabetizados, podem ter seu ensino facilitado através de ferramentas digitais que permitam a transcrição da fala. Atualmente sabe-se que este público é de aproximadamente 278 milhões de pessoas em todo o globo, sendo em média 5% da população dos países (PORTELLA et al., 2021). O uso de legendas como recurso pedagógico permite ao aluno surdo e ao professor dispensarem o uso de Libras e também a visualização da face do professor para a leitura labial. Desta forma, esta tecnologia, quando devidamente aplicada como ferramenta pedagógica, pode tornar-se relevante para a aprendizagem de conteúdo por parte desses educandos.

A contribuição para uma maior inclusão seria se as escolas adotassem e fizessem uso dessa tecnologia de transcrição para legendar todas as aulas virtuais, bem como as aulas presenciais simultaneamente, oferecendo dessa forma mecanismos que possibilitem a construção de uma aprendizagem significativa, sem distinção decorrente da limitação de alguns discentes, em especial, aos alunos da Educação Profissional e Tecnológica - EPT, objeto deste estudo.

Este artigo teve como objetivo descrever as ferramentas de legendas simultâneas disponíveis por meio de *software* livre e/ou ambiente web e que podem ser utilizados no ambiente escolar, visando melhorar a aprendizagem de alunos com deficiência auditiva.

MATERIAIS E MÉTODOS

Metodologicamente, este artigo pautou-se numa revisão bibliográfica exploratória, sistemática e qualitativa, cujo foco de pesquisa se dedicou a investigar a ocorrência de *softwares* e *sites* capazes de promover legendas simultâneas durante uma conversação.

As fontes utilizadas para a pesquisa bibliográfica foram artigos disponíveis na *Scientific Eletronic Library Online* (SciELO), teses, dissertações, *sites* institucionais e *blogs* que tratam da promoção de tecnologias que podem ser utilizadas como ferramentas de legendas simultâneas.

O protocolo de busca bibliográfica foi executado da seguinte maneira: a) limitou-se o período de publicação das fontes utilizadas aos anos de 2017 a 2022, b) os materiais bibliográficos foram selecionados através de leituras críticas e analíticas, c) após a análise das fontes consultadas, recorreu-se a discussão dos dados para a fundamentação dos argumentos (MARCONI; LAKATOS, 2011).

RESULTADOS E DISCUSSÃO

A empresa multinacional americana de serviços *online* e *softwares* Google, que hospeda e distribui diversos produtos, tem se destacado por seu protagonismo no que se refere ao desenvolvimento de tecnologias de transcrição da fala, apresentando para o público em geral diferentes aplicativos. A oferta de apps diversos com esta finalidade se deve ao *Project Euphonia*, que busca construir canais de maior acessibilidade aos usuários de seus produtos (VICENT, 2019).

Dentre as tecnologias de transcrição da linguagem oral para a textual utilizadas em computadores, *notebook* e *tablets* destacam-se o *Live Transcriber*, o *Live Caption* e o *Live Relay*. Lançado em 2019, este aplicativo (app), que em tradução literal significa ‘tradução em tempo real’ é um app desenvolvido pelo pesquisador Dimitri Kanevsky e pelo engenheiro de *software* Chet Gney e disponibilizado pelo Google, que hospeda e distribui, mediante pagamento este produto (VICENT, 2019).

Atualmente, o uso de Tecnologias Assistivas (TA) têm sido uma das principais ferramentas utilizadas pelas pessoas com deficiência auditiva e Surdos Usuários da Língua Portuguesa (SULP) para a transcrição da linguagem oral em escrita, fato este que permite uma melhor mediação da linguagem falada. Sucintamente, Silva; Mendes e Santos (2020) descrevem as Tecnologias Assistivas como aquelas cuja principal funcionalidade é a exposição e interação de seu usuário, via tela de computador, com conteúdo ao qual está sendo exposto. Estes mesmos autores afirmam que “[...] a proposta da TA está em ampliar a comunicação, a mobilidade, o controle do ambiente, as possibilidades de aprendizado, trabalho e integração na vida familiar, com os amigos e na sociedade (SILVA; MENDES; SANTOS, 2020, p. 4).

Uma das primeiras Tecnologias Assistivas disponibilizadas para o público visando unir tecnologia e conversação foi o Gboard. Disponível para usuários de *smartphones* com sistema operacional Android ou IOS, o Gboard é um aplicativo desenvolvido pelo Google e lançado em 2016 para ser utilizado como teclado virtual. Neste caso, basta ao usuário utilizar o microfone disponível no app de mensagem, e, enquanto fala, o aplicativo faz o reconhecimento de voz e produz a mensagem de texto a ser enviada (FREITAS, 2020).

O Gboard tem sido bastante utilizado para a realização de diálogos com surdos usuários da língua portuguesa. Contudo, já existem no mercado alguns aplicativos com a oferta de maiores recursos para serem utilizados durante as conversações. A seguir serão descritas algumas das tecnologias existentes que permitem aos surdos usuários da língua portuguesa utilizarem mídias digitais através da exposição de textos. Contudo, é preciso destacar que, infelizmente, são raros os textos científicos que pautam esse tema, fato este que ocasionou o uso, na maior parte das vezes, de autores de *blogs* e *sites* de tecnologia e informática.

Live Transcriber

É importante destacar que o *Live Transcriber* surgiu por dois fatores distintos. Primeiramente, antes mesmo do Google viabilizar a criação deste aplicativo, Kanevsky, que é matemático por formação e surdo desde a infância, procurou desenvolver, já na fase inicial como colaborador desta empresa americana, um algoritmo que permitisse reproduzir uma técnica denominada estenotipia. Por intermédio desta, um indivíduo escreve em tempo real, através de uma máquina denominada estenógrafo, o que é dito em seções como interrogatórios e audiências (MOREIRA, 2022). Atualmente, de acordo com Shimosakai (2012), esta técnica é muito utilizada no serviço de *closed caption*, que pela Associação Brasileira de Normas Técnicas (ABNT) exige 98% de exatidão.

O algoritmo criado com a intenção de descrever a linguagem oral, ou seja, buscando identificar os sons emitidos durante o processo de fala foi o *Extended Baum-Welch*, que tinha a capacidade de realizar o reconhecimento de voz. Para isto, Kanevsky recorreu à matemática aplicada através dos Modelos de Markov Ocultos, considerados um dos métodos mais eficientes para se calcular a verossimilhança entre parâmetros. É importante frisar, contudo, que diferentes comunidades que lidam com a questão da tradução simultânea da fala buscam desenvolver estes algoritmos, recorrendo principalmente, para isto, a parâmetros de misturas gaussianas (KANEVSKY et al., 2008).

O *Live Transcriber* pode ser acessado por *smartphones*, estando disponível em mais de 80 línguas. No caso dos aparelhos com sistema operacional Android, o usuário

deverá acessar o *Google Play Store*, buscar por *Live Transcriber* e fazer o *download* do aplicativo. Por sua vez, para baixá-lo no sistema operacional Apple, deve-se acessar o app *Live Transcriber* através do *App Store* (RIBEIRO, 2019). Após a sua instalação, em ambos os sistemas, seu uso é simples, pois ele mostra na tela do celular, sob a forma de texto, todo o conteúdo que foi falado.

Numa situação de conversação, o surdo usuário da língua portuguesa deve ativar o app no *smartphone* e colocá-lo ao alcance do indivíduo com quem conversa, para que seu aparelho possa captar com qualidade o som da voz da pessoa com a qual conversa. Assim, após esta captação, o usuário recolhe para si o aparelho e, olhando a tela, faz a leitura, dando prosseguimento à conversação.

Live Caption

Este aplicativo, assim como o *Live Transcriber*, foi desenvolvido pelo Google e utiliza o algoritmo *Extended Baum-Welch* para fazer o reconhecimento de voz e assim realizar a transcrição das conversações. Dessa forma, o *Live Caption* é descrito como um aplicativo capaz de realizar legendagem em tempo real para áudios disponíveis em aparelhos de *smartphone* que o tenham instalado. Para isto, o usuário deverá acessar o *Google Play Store* e baixar o app (TILMAN, 2021).

O uso do *Live Caption* não exige conexão com a *internet*, exceto para fazer o seu *download*, fato que facilita o acesso ao aplicativo em locais onde a rede de *Wi-Fi* ou de dados não seja efetiva. Uma vez instalado, para que seja utilizado, o usuário deve apenas acionar o app através do ícone de volume do aparelho, fazendo com que apareça a opção “usar legenda instantânea”. Assim, ao aproximar o aparelho do emissor durante a conversação, aparecerá no visor do *smartphone* a legenda do que foi falado (MOREIRA, 2020).

Além disso, o *Live Caption* tem uma particularidade interessante para as pessoas com deficiência auditiva e os SLP: o aplicativo permite o uso de legendas instantâneas durante a realização de chamadas. Neste caso, ao fazer uma ligação, a indicação de que a chamada será legendada aparece no visor do receptor, indicando para o mesmo que é necessário utilizar o teclado para atendê-la (MOREIRA, 2020).

No entanto, alguns aperfeiçoamentos ainda são necessários para que o *Live Caption* se torne ainda mais eficaz. Primeiramente, é importante destacar que ainda não há disponibilidade de legendagem em português durante o uso deste app. Outro fator que pode ser um complicador é que ainda não é possível salvar as conversações legendadas desse aplicativo, o que dificultaria o seu uso por um surdo usuário da língua portuguesa. Outra desvantagem deste aplicativo é que, para ser usado, ele utiliza uma considerável carga de energia, fazendo com que a bateria do

aparelho descarregue mais rápido (SUPPORT.GOOGLE.COM, 2022).

Live Relay

Conhecido em tradução livre como “retransmissão ao vivo”, o *Live Relay* surgiu da necessidade de aprimoramento de alguns algoritmos responsáveis por reconhecimento da voz, cujo uso foi destinado, primariamente, para indivíduos que possuem alguma enfermidade que impossibilite a comunicação de modo fluente. Assim, busca-se através do *Live Relay* captar a voz de pessoas com dificuldades na fala ocasionadas por doenças como esclerose lateral amiotrófica (ELA), lesões cerebrais traumáticas (LCT), Parkinson e acidente vascular cerebral (AVC) (SILVA, 2019).

Assim como o *Live Transcriber* e o *Live Caption*, o *Live Relay* utiliza o algoritmo *Extended Baum-Welch* para fazer o reconhecimento de voz e transcrever as conversações. Uma vez identificada a voz do usuário, ele é capaz de converter a fala em formato de texto ou, caso seja enviado a mensagem textual, o aplicativo transforma o texto em mensagem de voz. Para isto, os pesquisadores da empresa Google testaram através de experimentos o padrão de voz de pessoas com ELA de modo que o app fosse capaz de reproduzi-los através de um padrão visual reconhecível. Ao estabelecer esta relação, a inteligência artificial do *Live Relay* calibra este novo padrão de voz de modo que se permita fazer o seu reconhecimento tornando possível a transformação das palavras ditas em frases escritas. Além de converter áudio em texto, o *Live Relay* é capaz de transformar textos em áudio, fato este que facilita a conversação (MOREIRA, 2020).

Não obstante, ele apresenta como aspecto positivo o fato de que sugere frases ao usuário durante a digitação do texto, o que aumenta a velocidade da conversação, melhorando a experiência com o aplicativo. Embora seu uso seja direcionado às pessoas com dificuldade na fala, em decorrência de problemas de saúde, vê-se que o *Live Relay* pode ser facilmente utilizado por pessoas com deficiência auditiva e surdos usuários da língua portuguesa. Destaca-se, contudo, que este aplicativo está disponível para *download* apenas em sistemas operacionais Android e as legendas estão disponíveis em inglês (SILVA, 2019).

Web Captioner

Traduzido de modo literal como “legendador web”, o *Web Captioner* é um transcritor de áudio *online* que pode ser utilizado de modo gratuito pelo usuário. Disponível em 40 línguas e dialetos, o *Web Captioner* usa o sistema de reconhecimento por voz desenvolvido pelo Google e lançado em 2012. Desta forma, através do *Web Speech API*, foi possível aos desenvolvedores deste programa utilizarem o recurso de identificação de voz para que as conversações pudessem ser transcritas num navegador web (CERVO, 2021).

O principal atributo que torna o *Web Captioner* um transcritor que pode ser utilizado de modo massivo em conversações com surdos usuários da língua portuguesa é o fato de ser gratuito. Embora já estejam disponíveis no mercado alguns aplicativos que realizam a transcrição de áudios, o acesso muitas vezes é dificultado em virtude do valor cobrado para seu *download*. Além disso, o *Web Captioner* possibilita: a) transcrição em tempo real; b) transcrição de vídeos; c) transcrição de gravações de áudio e, d) apresenta boa qualidade textual na transcrição de áudios (MOREIRA; 2021).

O uso dele é de fácil uso. Para isto, é necessário ao usuário acessar o *link* webcaptioner.com/captioner. Posteriormente, deve-se selecionar a língua desejada através do ícone “*settings*”, localizado na aba configurações no canto inferior da tela. Feito isso, retorna-se à tela inicial do programa. Ao optar por “*start captioning*”, o navegador irá solicitar ao usuário a permissão para utilizar o microfone do aparelho que estiver utilizando para a conversação. Ao permitir o uso do microfone e selecionar a sua melhor configuração, o programa já estará apto a captar as mensagens ditas pelo usuário, sendo interrompidas a qualquer momento através do ícone “*stop captioning*”. Caso a transcrição seja feita de um vídeo, deve-se ter um *drive* apropriado instalado no aparelho a ser utilizado. Por sua vez, se o desejo do usuário é apagar a transcrição, deve-se acessar o ícone de exclusão, representado por uma lixeira (CERVO, 2021).

Embora seja significativa a importância do *Web Captioner* para as pessoas com deficiência auditiva e aos SULP, alguns ajustes ainda são necessários. Dentre eles, destacam-se: a) a flexibilização de uso para diferentes navegadores, já que se utiliza somente o *Chrome*; b) o aprimoramento do uso da pontuação nas frases transcritas; c) uso no modo *offline*; d) a substituição de instalação de *drive* por outra ferramenta já disponível nos aparelhos utilizados para transcrição quando o usuário buscar transcrever um vídeo; e) ausência de versão para o *Linux* (CERVO, 2021).

É importante destacar que o *Web Captioner* atende satisfatoriamente às pessoas com deficiência auditiva e surdos usuários da língua portuguesa quando utilizados em sala de aula quando associado a algum equipamento de projeção. Assim, é possível obter o uso do *Web Captioner* e o aparelho de projeção conectado a um computador concomitantemente, a partir de uma gravação onde será feita a transcrição e posterior projeção, e também associado a uma apresentação em *powerpoint*, onde o microfone do computador do professor pode ser ligado e a legendagem ser produzida na projeção juntamente com o material a ser exposto (MOREIRA; 2021).

CONCLUSÃO

O uso da tecnologia como facilitadora da interação entre os surdos usuários da língua portuguesa, principalmente nos diálogos que se estabelecem em sala de aula, destacam a

relevância deste trabalho por descrever diferentes ferramentas de legendas simultâneas disponíveis por meio de *software* livre e/ou ambiente *web*.

A carência de publicações relacionando o uso de legendas simultâneas para os usuários com deficiência auditiva e SULP evidencia a necessidade de enfatizar a inclusão escolar destes alunos sob diferentes variáveis culturais. Nesta perspectiva, este artigo pode ser considerado de grande valor já que se viabiliza como ponto de partida para futuras buscas bibliográficas sobre a inclusão digital e escolar de alunos surdos.

As principais ferramentas de tradução simultânea encontradas durante a pesquisa apontaram a busca pela inclusão digital de leitores usuários com deficiência auditiva e surdos por parte de seus desenvolvedores. O incentivo à conversação, através da linguagem oral mediada pela linguagem escrita, facilita a interação social entre os indivíduos, principalmente no ambiente escolar. A elaboração de legendas simultâneas somente foi possível após o desenvolvimento do projeto *Euphonia*, desenvolvido por Dimitri Kanevsky e pelo engenheiro de *software* Chet Gney. Destaca-se, contudo, que o fato de Dimitri ser surdo foi ponto crucial para o desenvolvimento do projeto e pela elaboração dos programas.

Por fim, este artigo desponta como mais um suporte necessário a futuros trabalhos, ao permitir uma reflexão sobre as potencialidades que as legendas simultâneas podem ocasionar às pessoas com deficiência auditiva e surdos usuários da língua portuguesa. Sua relevância se destaca pela metodologia adotada, que converte os dados coletados em informações depuradas pelo rigor científico.

REFERÊNCIAS

1. PORTELLA, S. M.; GOUDINHO, L. DA S.; FERREIRA, A. T. S.; MENDES, M. C. B.; VALE, M. R. M. DOS S.; OLIVEIRA, A. F. DE; LEITE, E. A.; SILVA JÚNIOR, E. DOS S.; FAUSTO, I. R. DE S.; PINTO, S. C. V. DA S.; BRAZ, R. M. M. As bases biológicas da surdez. **Research, Society and Development**, v. 10, n. 10, 2021
2. MARCONI, M. DE A.; LAKATOS, E. M. **Metodologia do trabalho científico**: procedimentos básicos, pesquisa bibliográfica, projeto e relatório, publicações e trabalhos científicos. 7. ed. 6.reimpr. São Paulo: Atlas, 2011.
3. VICENT, J. **O Projeto Euphonia do Google ajuda a tornar a tecnologia de fala mais acessível a pessoas com deficiência**. 2019. Disponível em: <<https://www.theverge.com/2019/5/7/18535674/google-project-euphonia-live-relay-speech-recognition-disabilities-impairments-io-2019>>. Acesso em: 03 jul. 2022.
4. SILVA, Q. P. DA; MENDES, N. F. O.; SANTOS, S. K. DA S. L. Tecnologia assistiva como fator de

- acessibilidade no processo de ensino aprendizagem de surdos. Congresso Internacional de Educação e Tecnologia. **Anais do CIET:EnPED:2020 - (Congresso Internacional de Educação e Tecnologias | Encontro de Pesquisadores em Educação a Distância)**, São Carlos, ago. 2020. ISSN 2316-8722. Disponível em: <<https://cietenped.ufscar.br/submissao/index.php/2020/article/view/1858>>. Acesso em: 03 jul. 2022.
5. FREITAS, F. **5 funções pouco conhecidas do Gboard que você precisa conhecer**. 2020. Disponível em: <<https://canaltech.com.br/apps/gboard-melhores-funcoes-dicas/>>. Acesso em: 03 jul. 2022.
6. MOREIRA, P. **10 perguntas para o inventor do *google live transcribe*: Dimitri Kanevsky**. 2022. Disponível em: <<https://cronicasdasurdez.com/dimitri-kavensky-google-live-transcribe/>>. Acesso em: 03 jul. 2022.
7. SHIMOSAKAI, R. **Conheça a profissão do Estenotipista, responsável pela Legenda Oculta**. 2012. Disponível em: <<https://ricardoshimosakai.com.br/>>. Acesso em: 03 jul. 2022.
8. KANEVSKY, D.; SAINATH, T. N.; RAMABHADRAN, B.; NAHAMOO, D. ***Generalization of Extended Baum-Welch Parameter Estimation for Discriminative Training and Decoding***. 2008. Disponível em: <https://groups.csail.mit.edu/sls/publications/2008/Sainath2_Interspeech08.pdf>. Acesso em: 03 jul. 2022.
9. RIBEIRO, C. **Google lança apps de acessibilidade *Live Transcribe* e Amplificador de Som**. 2019. Disponível em: <<https://www.techtudo.com.br/noticias/2019/02/google-lanca-apps-de-acessibilidade-live-transcribe-e-amplificador-de-som.ghtml>>. Acesso em: 03 jul. 2022.
10. TILLMAN, M. **O que é *Google Live Caption* e como funciona?** 2021. Disponível em: <<https://www.pocket-lint.com/pt-br/celulares/noticias/google/149794-o-que-e-legenda-ao-vivo-do-google-e-como-funciona>>. Acesso em: 03 jul. 2022.
11. SUPPORT.GOOGLE. **Converter voz em texto com a Legenda instantânea**. 2022. Disponível em: <<https://support.google.com/accessibility/android/answer/9350862?hl=pt-BR>>. Acesso em: 03 jul. 2022.
12. SILVA, V. H. **Google leva reconhecimento de voz a quem tem dificuldade de fala**. 2019. Disponível em: <<https://tecnoblog.net/noticias/2019/05/07/google-reconhecimento-voz-dificuldade-fala-io19/#:~:text=O%20reconhecimento%20de%20voz%20pode,acabam%20tendo%20seu%20uso%20limitado>>. Acesso em: 03 jul. 2022.
13. CERVO, G. C. **Transcrição de Áudio com *Web Captioner* para Windows 10**. 2021. Disponível em: <<https://www.ufsm.br/app/uploads/sites/762/2021/02/Tutorial-Transcricao-de-Audio-com-Web-Captioner.pdf>>. Acesso em: 03 jul. 2022.
14. MOREIRA, P. **Tutorial para legendar com o *Web Captioner***: tenha legenda em tudo, sempre. 2021. Disponível em: <<https://cronicasdasurdez.com/tutorial-wecaptioner-legendas/>>. Acesso em: 03 jul. 2022.