

Avaliação da Utilização da Conversão de Fala para a Escrita para Melhoria da Qualidade da Comunicação dos Deficientes Auditivos

Henry Souza Cândido, Marcelo C. P. Santos

Instituto de Informática – Instituto Federal de Educação Ciência e Tecnologia do Sudeste de Minas Gerais (IFSudesteMG)

Juiz de Fora – MG – Brazil

{henry@hotmail.com,marcelocpsantos@gmail.com}

Abstract. *The objective of this work is to evaluate the software for converting speech into text described in scientific literature or available for installation in smartphones. We hope, therefore, to contribute to the effective use of these tools to improve the quality of life of the deaf. Two metrics will be used to evaluate conversions for text. The first is widely used in similar articles, and the second, a methodology proposed in this work that presents advantages over traditional metrics.*

Resumo *O objetivo desse trabalho é avaliar os “softwares” para conversão de fala em texto, descritos na literatura científica ou disponíveis para instalação em “smartfones”. Esperamos, dessa forma, contribuir para a efetiva utilização dessas ferramentas para melhoria da qualidade de vida dos surdos. Serão utilizadas duas métricas para das conversões. A primeira amplamente utilizada em artigos semelhantes, e a segunda, uma metodologia proposta nesse trabalho que apresenta vantagens sobre a métrica tradicional.*

1. Introdução

Segundo o último censo do IBGE de 2010, cerca de 9,8 milhões de brasileiros possuem deficiência auditiva, o que representa 5,2% da população brasileira. Cada vez mais essa parte da população vem buscando integração à sociedade e, naturalmente, acaba encontrando muita dificuldade durante o processo de comunicação e de aprendizado em geral. O uso de intérpretes em sala de aula tem ajudado bastante, contudo pensando num maior grau de independência do deficiente, o uso de “softwares” de conversão de fala em texto pode ajudar ainda mais na educação destes indivíduos.

Com o desenvolvimento dos “softwares” que convertem fala para texto (STT – “speech to text”), hoje utilizados corriqueiramente, acreditamos ser possível melhorar consideravelmente a comunicação com deficientes auditivos em aulas, palestras, conversas telefônicas, e no dia a dia de forma geral. Os surdos que utilizam LIBRAS no dia a dia em lugar do português oral, normalmente têm dificuldades no português escrito, o que amplia o desafio de sua inclusão, principalmente no meio acadêmico onde grande parte do conhecimento encontra-se disponível na forma escrita e trabalhos e provas são tradicionalmente entregues nessa forma de expressão. Com o uso do STT, o surdo tende a melhorar seu entendimento do português escrito e podemos esperar que

consequentemente melhore também suas habilidades para expressar com essa modalidade.

Observamos também que os gestos de LIBRAS são construídos em função da necessidade de inclusão de novos termos ao vocabulário corriqueiro de seus usuários. Quando não existe um gesto criado, o termo é soletrado, o que dificulta bastante a comunicação. Quando um surdo ingressa em um meio acadêmico, todo um novo vocabulário lhe é apresentado. Até que um vocabulário LIBRAS seja construído e difundido, acreditamos que o português e o STT podem ser de grande importância na acessibilidade do surdo à conteúdos e jargões de áreas específicas do conhecimento.

Nesse trabalho levantamos os principais “softwares” que implementam o STT citados em artigos científicos e encontrados nas duas principais lojas para aplicativos para smartphones ("Google play" e "Apple Store"). Avaliamos a qualidade da tradução segundo dois critérios: o WER - Word Error Rate e o CCP - Critério de Compreensão de Palavras, o último proposto nesse trabalho. Além disso, avaliamos a facilidade de uso e instalação, qualidade da documentação.

2. Levantamento Bibliográfico

Primeiramente foi realizada uma pesquisa por artigos científicos nas principais bases acadêmicas. As fontes da busca foram: Capes, Google Acadêmico, “ACM Digital Library”, “IEEE Digital Library”, “Science Direct”, Scopus. Foram procurados artigos escritos em Português e Inglês, publicados após 2015.

Utilizou-se a string de busca:

(“surdo” OR “deaf”) AND

("speech to text" OR “stt” OR "tecnologia assistiva") OR

(“conversão” AND “voz” AND “texto”)

-("deficiência visual")

Foram encontrados 47 artigos. Após a leitura de títulos e resumos detectou-se que 35 desses artigos não eram de interesse central para o trabalho, focando principalmente em aspectos sociais da educação para surdos bem como políticas públicas ou abordagens pedagógicas na educação de surdos.

Furlan (2016) analisa vários tipos de tecnologias assistivas de forma muito semelhante à proposta desse trabalho. O artigo propõe o desenvolvimento de um aplicativo, o “Assistive Messenger”, mas não existe versão disponível do “software” proposto para testes; Grimmes (2017) também faz uma análise comparativa de aplicativos que fazem tradução para surdos, no entanto, os “softwares” analisados no artigo não se encontram mais disponíveis para teste; Santos e Dantas (2017) também fazem uma exposição de “softwares” (“Handtalk”, “Prodeaf” e “Rybena”) e outros “softwares” de LIBRAS, mas, no entanto não propoem uma comparação entre eles.

Alain e Verjano (2016) tratam da situação da inclusão da comunidade surda do Panamá, sem explorar tecnologias;

Soluções que limitam a comunicação a frases pré-determinadas e situações específicas do cotidiano são descritos em Carneiro (2016) e Jhon (2016) sendo que o

último propõe o desenvolvimento de um equipamento dedicado com “software” embarcado.

Pivetta, Ulbricht e Savi (2011), fazem uma análise comparativa de alguns tradutores de libras; Furui (2007) propõe duas formas de se avaliar os “softwares” de reconhecimento de voz e de falantes para o japonês, seguindo critérios objetivos e subjetivos, o trabalho propõe uma metodologia inovadora para avaliação de traduções em Japonês. Stuker, Kilgour e Niehues (2012), expõem um projeto de avaliação de tecnologias assistivas conhecido com Quaero sem, no entanto, propor uma metodologia própria, apenas apresentando as características do projeto. Venkateswarlu, Raviteja e Rajeev (2012) descrevem todo processamento da palavra pelos “softwares” de reconhecimento de voz, e propõem uma análise objetiva de cada fase, através de fórmulas matemáticas para estudar se o processamento e reconhecimento da fala estão de acordo com o esperado, e uma análise comparativa de vários usuários testados. Nobuyasu (2014) propõe um método que mede a taxa de erro de caracteres em uma palavra, padrão que seria mais adequado ao idioma japonês.

E por fim Martins Santo (2017) fornece uma visão geral dos três reconhecimentos de voz automáticos: Vocapia VoxSigma™, “Google Cloud Speech”, e Limecraft Transcribe. O reconhecimento é usado em sistemas de STT, e no caso ele o usou em gravações de áudio históricas, aplicando a taxa de erro de palavra (WER) para avaliá-los.

Dentre os aplicativos mencionados na literatura selecionamos os de interesse (ver tabela 1) que incluam em suas funções STT (“Speech To Text”), embora também possam ter outras funções sendo eles: “Assistive Messenger”, “Scribe”, Enssat, “Handtalk” e “Prodeaf”.

Tabela 1. Relata os “softwares” mencionados nos artigos e sua avaliação.

“software”	Artigo	Avaliação
“Assistive Menssenger”	Furlan (2016)	Descartado: Protótipo apenas.
“Scribe”	Lasecki (2017)	Descartado pois não faz STT automático, é um “software” que auxilia o homem para fazer a transcrição
“Enssat”	Santos e Dantas (2017)	Descartado: não disponível
“Handtalk”	Furlan (2016), Grimmes (2017), Santos e Dantas (2017)	Avaliado
“Podeaf”	Furlan (2016), Grimmes (2017)	Avaliado

3. “Software” para “Smartfones”

Na “Apple Store”, foram encontrados e avaliados os seguintes “softwares” gratuitos:

- “Speech to text dictate”,
- “Speechy”,
- “Dictate”,
- “Handtalk”,
- “Prodeaf”.

No “Google Play” selecionamos aplicativos gratuitos, avaliados pelos usuários com quatro ou cinco estrelas, e última atualização em 2017 ou 2018. Foram então avaliados:

- “Speech to text/text to speech”,
- “Speechnotes”,
- “Speech to text translator tts”,
- “Listnote fala para texto notas”,
- “G-board”.

4. Critérios de Avaliação

Avaliamos os “softwares” segundo os seguintes quesitos:

- Qualidade da Tradução;
- Instalação;
- Usabilidade;
- Documentação;

Para os três últimos quesitos, atribuiu-se um conceito como segue:

1. Insatisfatório;
2. Ruim;
3. Regular;
4. Ótimo;
5. Excelente;

4.1 Avaliação da Qualidade da Tradução

Para a avaliação da qualidade da tradução utilizou-se duas métricas: A Taxa de Erros de Palavras (WER – “Word Error Rate”) e o CCP (Critério de Compreensão de Palavras).

O WER é uma métrica frequentemente utilizada para avaliação de sistemas de reconhecimento de fala ou de tradução automática. Exemplos de seu emprego podem ser observados em Nobuyasu (2014) e Martins Santo (2017). Segundo estes, neste tipo de sistema três erros são considerados comuns: substituição de palavras (quando uma palavra dita corretamente é traduzida de maneira diferente); Exclusão (quando uma palavra dita corretamente não é transcrita); inserção (quando uma palavra extra, que não foi pronunciada é inserida na transcrição).

A taxa de erro de palavras pode então ser calculada como:

$$WER = (S+D+I) / (S+D+C);$$

Onde:

S é o número de substituições,

D é o número de exclusões,

I é o número de inserções, e

C é o número de palavras corretas

O WER é derivado da distância de Levenshtein, trabalhando no nível de palavra em vez do nível de fonema sendo uma ferramenta valiosa para comparar sistemas diferentes, bem como avaliar melhorias em um sistema. Este tipo de medição, no

entanto, não fornece detalhes sobre a natureza dos erros de tradução, por isso também propusemos um método complementar de avaliação.

O WER atribui o mesmo peso a qualquer erro de tradução, no entanto, observamos que existem erros que não trazem qualquer prejuízo à compreensão da frase, enquanto outros tornam a mensagem transmitida completamente inválida. Propomos então o CCP, Critério de Compreensão de Palavras, que tem por objetivo, além de medir o número de erros, avaliar o quanto esses erros comprometem a compreensão da mensagem, objetivo mais importante da comunicação.

- Erro de peso 10: Compromete entendimento da frase. Exemplo: Os “super” subiram ontem (preços).
- Erro de peso 5: Apesar do erro, a palavra pode ser compreendida em função do contexto da frase. Se retirarmos a palavra do contexto (da frase) a ela se torna ininteligível. Exemplo: Os preços “biram” ontem (subiram).
- Erro de peso 1: Apesar do erro, a palavra pode ser entendida mesmo quando analisada fora da frase. Exemplo: Preciso ir ‘o’ médico (ao).

Para ambos os métodos as frases da lista que segue foram apresentadas de forma oral (falada) para o “software” e avaliou-se o texto gerado. Utilizamos sempre o mesmo locutor para todos os “softwares”. A lista foi extraída de Barreto (2010). O autor utilizou a mesma lista em um estudo fonoaudiológico sobre inteligibilidade do discurso oral, e alega em seu trabalho que as frases possuem distribuição semelhante de fonemas, estruturas silábicas e extensão das palavras, o que torna a lista desenvolvida adequada também para nosso estudo. O conjunto de frases é:

1. Não posso perder o ônibus.
2. Vamos tomar um cafezinho.
3. Preciso ir ao médico.
4. A porta da frente está aberta.
5. A comida tinha muito sal.
6. Cheguei atrasado para a reunião.
7. Vamos conversar lá na sala.
8. Depois liga pra mim.
9. Esqueci de pagar a conta.
10. Os preços subiram ontem.
11. O jantar está na mesa.
12. As crianças estão brincando.
13. Choveu muito nesse fim de semana.
14. Estou morrendo de saudade.
15. Olhe bem ao atravessar a rua.
16. Preciso pensar com calma.
17. Guardei o livro na primeira gaveta.
18. Hoje é meu dia de sorte.
19. O sol está muito quente.
20. Sua mão acabou de sair de carro.
21. Ela vai viajar nas férias.
22. Não quero perder o avião.
23. Eu não conheci sua filha.
24. Ela precisa esperar na fila.

25. O banco fechou sua conta.

5. Análise dos Resultados

5.1 “Dictate”

Este “software” está presente apenas na Apple-Store, e faz conversão simultânea de voz para texto. Ele apresenta uma velocidade de conversão satisfatória e ágil, porém não permite visualizar o texto por inteiro devido à sua fonte extremamente grande, e uma caixa de texto pequena. Está presente na forma de aplicativo móvel, gratuito (Figura 1).

5.1.2 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas da “Apple-Store”, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: Excelente (5).

5.1.3 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Porém seu uso é prejudicado pela fonte exagerada e pela caixa de texto pequena, o que impede de visualizar o texto por completo. Não apresenta pausas e propagandas.

Portanto a avaliação do mesmo seria considerada: INSATISFATÓRIO (1).

5.1.4 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória e simultânea para texto, com boa velocidade e poucos erros de tradução.

WER=0,024; S= 2; D=1; I=0; C=122.

CCP= 17. Deleção: (ao/o +1). Substituição: preços/super (+10); subiram/biram (+5); (o/do (+1)).

5.1.5 Avaliação de Documentação

O “software” conta com descrição através da Apple Store. Nele pode-se encontrar um material descritivo do produto, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

Desta forma sua avaliação para a documentação foi considerada: OTIMO (4).

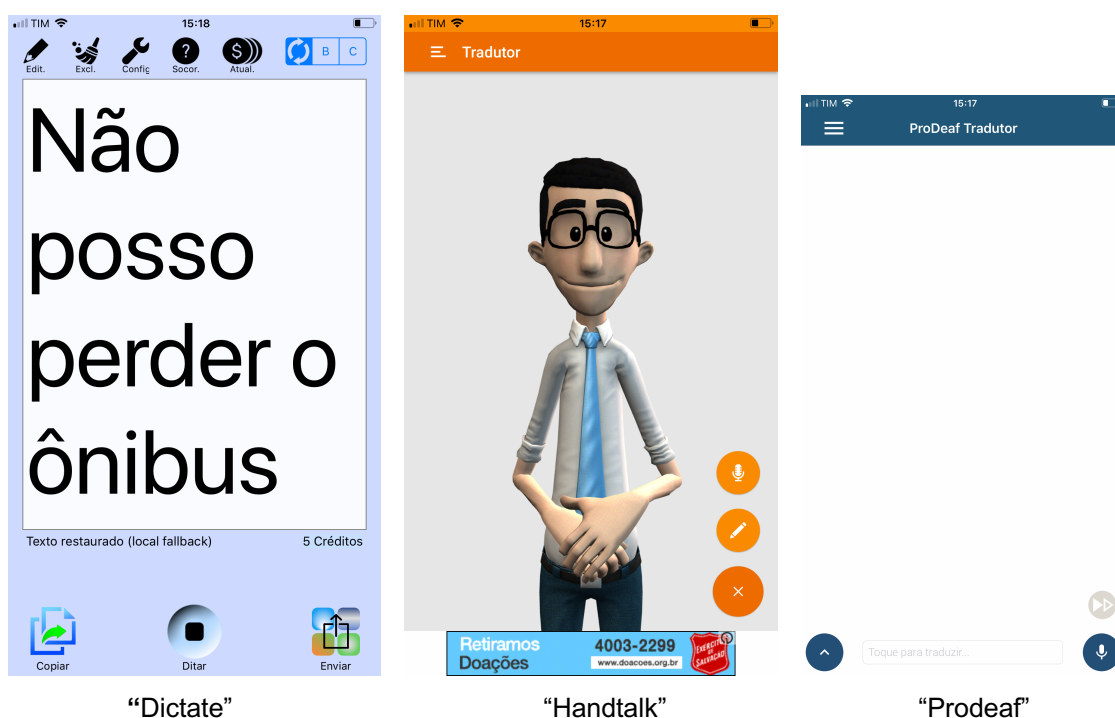


Figura.1 Observamos as telas dos “softwares” “Dictate” (seção 5.1), “Handtalk” (seção 5.2) e “Prodeaf” (seção 5.3).

5.2 “Hand Talk”

Este “software” foi desenvolvido por uma equipe de brasileiros, e se assemelha muito ao projeto do Prodeaf. Sendo que em 2013, foram premiados como o Melhor Aplicativo Social do Mundo pela ONU. Em junho de 2018 a “HandTalk” comprou a companhia produtora do “ProDeaf”.

Possui a tradução simultânea de voz para Libras e para texto (Figura 1), e também de texto para som ou libras. Sendo que a velocidade de tradução pode ser controlada por um velocímetro, contudo é de fato mais lenta que a do “Prodeaf”, o que dificulta sua utilização. Sua interface é organizada de forma simples e intuitiva, semelhante ao “ProDeaf”.

No caso de Libras o “software” também conta com um “avatar” chamado Hugo que realiza os movimentos das mãos como se fosse um intérprete real. Nele também existe a opção de tradução para Língua Americana de sinais. Sendo este um dos “softwares” mais usados pelos deficientes atualmente.

O “software” está disponível em versões para web, IOS e Android e uma versão para existe para empresas que desejem ter suas páginas na internet traduzidas para libras.

5.2.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas da “Apple Store” e “Google Play”, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: Excelente (5).

5.2.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Seu único porém se refere à velocidade de tradução, um pouco mais lenta. Prejudicada pelas pausas, e também pelas propagandas na versão gratuita, apesar de ter poucos erros de tradução, torna difícil a tradução em tempo real.

Portanto a avaliação do mesmo seria considerada: REGULAR (3).

5.2.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória, tanto para texto quanto para libras.

WER=0,04; S=3; D=2; I=0; C=120.

CCP=31. Substituições (sorte/sol +10, ao/ou +5, mesa/minhas +5); deleção (um ‘a’+1 e uma palavra sal +10).

5.2.4 Avaliação de Documentação

O “software” conta com ampla descrição e documentação através do site oficial da “Handtalk”. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas. Um grande diferencial é que a página da “HandTalk”, conta com o recurso de tradução para Libras, facilitando a navegação do deficiente auditivo, além de várias informações sobre a comunidade surda, de interesse da população ouvinte.

Desta forma sua avaliação para a documentação foi considerada: EXCELENTE (5).

5.3 “Prodeaf”

Este “software” (Figura 1) foi desenvolvido por uma equipe de brasileiros, ligados à Universidade Federal de Pernambuco.

Implementa a possibilidade de tradução simultânea de voz para Libras e para texto, sendo que na falta de internet, a tradução é feita apenas para texto. Com uma interface limpa e bem simples, torna sua utilização e navegação bem ágil.

No caso de Libras o “software” conta com um “avatar” que realiza os movimentos das mãos como se fosse um intérprete real. Nele também existe a opção de tradução para Língua Americana de sinais. Sendo este um dos “softwares” mais usados pelos deficientes atualmente.

Existe na versão móvel, como aplicativo, disponível para IOS e Android. Na versão Web, auxiliando na navegação pela internet, ambas gratuitas. E ainda numa versão paga, para empresas que queiram traduzir suas páginas automaticamente para Libras.

Há ainda o “Prodeaf QR Code”, capaz de ler e traduzir para libras códigos de produtos, contendo as informações nele descritas para a língua de sinais.

5.3.1 Avaliação de Instalação:

Como se trata de um aplicativo utilizado principalmente via Web, sua instalação é fácil, sendo encontrado gratuitamente nas lojas da “Apple Store” e “Google Play”, basta buscar e instalar diretamente no celular. Na sua versão web, também é possível baixá-la de forma rápida diretamente no site da empresa, que conta com informações úteis aos usuários e empresas.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: Excelente (5).

5.3.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação.

Portanto a avaliação do mesmo seria considerada: EXCELENTE (5).

5.3.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória, tanto para texto quanto para libras. Apresentou desempenho adequado para o texto traduzido de forma corrente, adaptável para o ambiente educacional. Não apresenta pausas, nem progandas. E com velocidade rápida de conversão.

WER=0,024; S=3; D=0; I=0; C=122.

CCP=15. Substituição (aberta/sabetta (+5), ao/ou (+5), quente/quendi (+5)).

5.3.4 Avaliação de Documentação:

O “software” conta com ampla descrição e documentação através do site oficial da Prodeaf. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas. Um grande diferencial é que a página da Prodeaf, conta com o recurso de tradução para Libras, facilitando a navegação do deficiente auditivo.

Desta forma sua avaliação para a documentação foi considerada: EXCELENTE (5).

5.4 “Speech to Text Dictate”

Este “software” está presente na “Apple Store”, e faz conversão simultânea de voz para texto. Ele apresenta uma velocidade de conversão satisfatória. Está presente na forma de aplicativo móvel, gratuito (Figura 2).

5.4.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente na loja da “Apple Store”, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: Excelente (5).

5.4.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Não apresenta pausas e propagandas.

Portanto a avaliação do mesmo seria considerada: EXCELENTE (5).

5.4.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória para texto. Poucos erros de tradução, e tradução simultânea satisfatória.

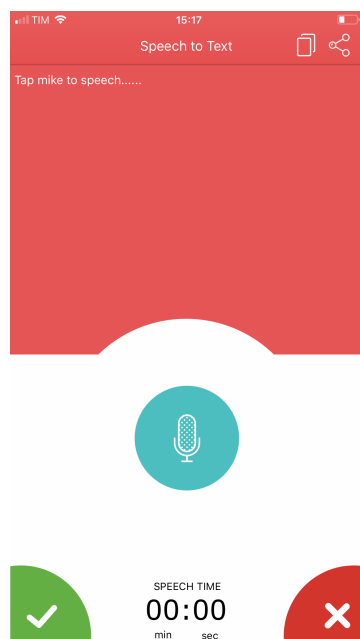
WER=0,008; S=1; D=0; I=0; C=124.

CCP= 1. Apenas trocou um ‘ao’ por ‘o’ (+ 1).

5.4.4 Avaliação de Documentação

O “software” conta com descrição através do “Google-play” e da “Apple Store”. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

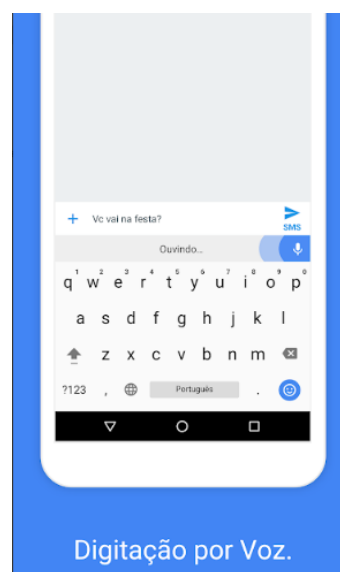
Desta forma sua avaliação para a documentação foi considerada: OTIMO (4).



“Speech toText”



“Speechy“



“Google Board”

Figura 2. Temos as telas do “Speech toText” (seção 5.4), “Speechy” (seção 5.5) e “Google Board” (seção 5.6).

5.5 “Speechy”

Este “software” está presente apenas na “Apple-Store”, e faz conversão simultânea de voz para texto. Ele apresenta uma velocidade de conversão satisfatória, bem rápida. Está presente na forma de aplicativo móvel, gratuito (Figura 2).

5.5.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas da “Apple-Store”, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: Excelente (5).

5.5.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Não apresenta pausas e propagandas.

Portanto a avaliação do mesmo seria considerada: EXCELENTE (5).

5.5.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória e simultânea para texto, com boa velocidade e poucos erros de tradução.

WER=0,032; S=2; D=1; I=1; C=122

CCP=7 (subst.: subiram/biram (+5); e ao/o (+1); deleção ‘o’ (+1); inserção ‘um’ (0).

5.5.4 Avaliação de Documentação

O “software” conta com descrição através da “Apple Store”. Nele pode-se encontrar um material descritivo do produto, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

Desta forma sua avaliação para a documentação foi considerada: OTIMO (4).

5.6 “Google Board”

Este “software” está presente na “Google-play”, e faz conversão simultânea de voz para texto. Ele apresenta uma velocidade de conversão satisfatória. Está presente na forma de aplicativo móvel, gratuito. Usa o teclado do Google para conversão (Figura 2).

5.6.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas “Google Play”, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: EXCELENTE (5).

5.6.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Sem pausas, traduz de forma rápida, contínua e instantânea. Usa o sistema Google de tradução.

Portanto a avaliação do mesmo seria considerada: EXCELENTE (5).

5.6.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória para texto. Poucos erros de tradução, e tradução simultânea insatisfatória, ocorre de forma muito lenta.

WER=0,016; S=2; D=0; I=0; C=124.

CCP= 20 (substituições: subiram/viram +10 e ao/eu +10).

5.6.4 Avaliação de Documentação:

O “software” conta com descrição através do “Google-play”. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

Desta forma sua avaliação para a documentação foi considerada: OTIMO (4).

5.7 Listnote Fala Para Texto Notas:

Este “software” está presente na “Google-play”, e faz conversão simultânea de voz para texto. Ele apresenta uma velocidade de conversão satisfatória. Está presente na forma de aplicativo móvel, gratuito (Figura 3).

5.7.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas “Google Play”, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: EXCELENTE (5).

5.7.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Sem pausas, traduz de forma rápida, contínua e instantânea.

Portanto a avaliação do mesmo seria considerada: EXCELENTE (5).

5.7.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória para texto. Poucos erros de tradução, e tradução simultânea insatisfatória.

WER=0,008; S=1; D=0; I=0; C=124

CCP= 10 (Olhe / ali + 10).

5.7.4 Avaliação de Documentação:

O “software” conta com descrição através do “Google-play”. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

Desta forma sua avaliação para a documentação foi considerada: OTIMO (4).

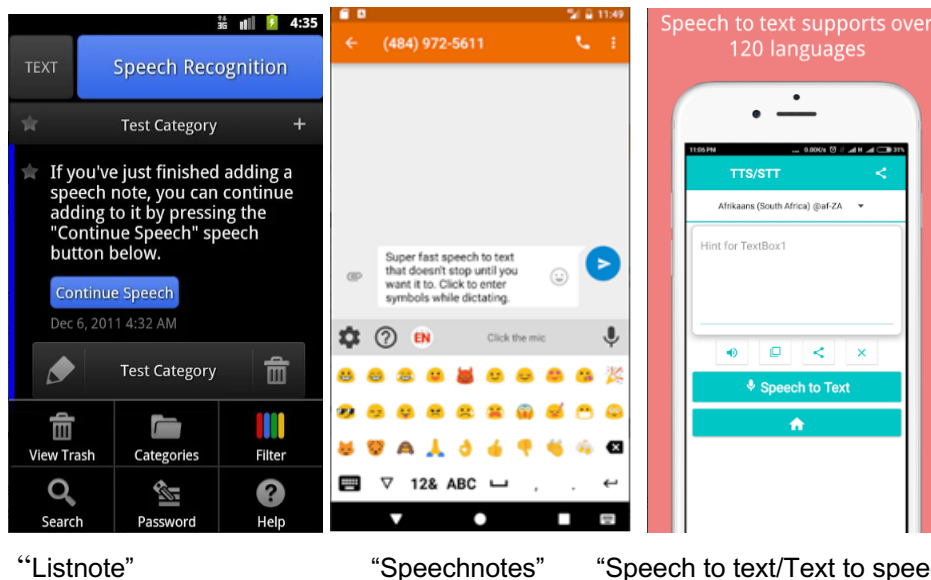


Figura. 3 Telas do “Listnote” Fala Para Texto (seção 5.7); “Speechnotes” (5.8); “Speech to text/Text to speech” (seção 5.9).

5.8 “Speechnotes”

Este “software” está presente na “Google-play”, e faz conversão simultânea de voz para texto. Ele apresenta uma velocidade de conversão satisfatória. Está presente na forma de aplicativo móvel, gratuito (Figura 3).

5.8.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas Google Play, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: EXCELENTE (5).

5.8.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Sem pausas, traduz de forma rápida, contínua e instantânea.

Portanto a avaliação do mesmo seria considerada: EXCELENTE (5).

5.8.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória para texto. Poucos erros de tradução, e tradução simultânea insatisfatória.

WER=0,008; S=1; D=0; I=0; C=124

CCP=5 (Substituição 'ao' por 'eu' (+5)).

5.8.4 Avaliação de Documentação:

O “software” conta com descrição através do “Google-play”. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

Desta forma sua avaliação para a transcrição foi considerada: OTIMO (4).

5.9 “Speech To Text / Text To Speech”

Este “software” está presente na “Google-play”, e faz conversão simultânea de voz para texto, porém não contínua. Ele apresenta uma velocidade de conversão satisfatória. Está presente na forma de aplicativo móvel, gratuito (Figura 3). Usa o sistema de reconhecimento de fala do Google.

5.9.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas “Google Play”, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: Excelente (5).

5.9.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Apresenta pausas e propagandas, e não faz a tradução de forma contínua.

Portanto a avaliação do mesmo seria considerada: RUIM (2).

5.9.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória para texto. Poucos erros de tradução, e tradução simultânea insatisfatória.

WER=0,008; S=1; D=0; I=0; C=12

CCP=5 (Substituição 'ao' por 'eu' (+5)).

5.9.4 Avaliação de Documentação:

O “software” conta com descrição através do Google-play. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

Desta forma sua avaliação para a documentação foi considerada: OTIMO (4).

5.10 Speech To Text Translator (TTS):

Este “software” está presente na Google-play, e faz conversão simultânea de voz para texto. Ele apresenta uma velocidade de conversão satisfatória. Está presente na forma de aplicativo móvel, gratuito (Figura 4). Usa o sistema Google de tradução.



Figura 4: Speech to text Translator.

5.10.1 Avaliação de Instalação:

Como se trata de um aplicativo móvel, sua instalação é fácil, sendo encontrado gratuitamente nas lojas Google Play, basta buscar e instalar diretamente no celular.

Seguindo esta linha de raciocínio, quanto à facilidade de instalação ele pode ser considerado: EXCELENTE (5).

5.10.2 Avaliação de Usabilidade:

O “software” possui uma estrutura intuitiva, organizada de forma simples e coesa.

Isto facilita muito na hora do uso, pois permite um acesso rápido aos recursos, requisito muito importante na hora do processo de comunicação. Sem pausas, traduz de forma rápida, contínua e instantânea. Usa o sistema Google de tradução.

Portanto a avaliação do mesmo seria considerada: EXCELENTE (5).

5.10.3 Avaliação de Tradução:

Como vimos o “software” faz a tradução de forma satisfatória para texto. Poucos erros de tradução, e tradução simultânea insatisfatória.

WER=0,008; S=1; D=0; I=0; C=124.

CPP=10 (Olhe / hoje (+ 10)).

5.10.4 Avaliação de Documentação:

O “software” conta com descrição através do Google-play. Nele pode-se encontrar um material descritivo dos produtos, sua instalação, uso e plataformas de ajuda para os usuários e empresas.

Desta forma sua avaliação para a documentação foi considerada: OTIMO (4).

6. Conclusão e trabalhos futuros

O critério para avaliação da qualidade de tradução proposto nesse artigo mostrou-se extremamente útil para a avaliação dos “softwares”. Se analisarmos os dois critérios, WER e CCP, verificamos que as duas grandezas apresentam alta correlação (0,71 para o coeficiente de Pearson), no entanto podemos observar que para alguns casos os valores apresentam diferenças grandes. O “Speechy” apresenta o pior índice WER (0,032), ou seja, foi o “software” que apresentou maior quantidade de erros, no entanto, o tipo de erro apresentado não foi relevante na compreensão das mensagens. O “Speechy” conseguiu índice 7 no CCP, quarto melhor entre os analisados.

A fim de elegermos o melhor “software” para STT entre os analisados, eliminamos os “software”s que obtiveram nota inferior a 4 em qualquer dos quesitos subjetivos, descartou-se, portanto o, “Speech To Text / Text To Speech” e o “Dictate”, que obtiveram notas 2 e 1 para usabilidade. Entre os oito “softwares” restantes, houve um empate com quatro sistemas apresentando índice WER igual a 0,008, no entanto, o “Speech to Text Dictate” obteve apenas 1 para o CCP, melhor do que todos os demais, portanto o CCP mostrou-se decisivo para a escolha do “Speech to Text Dictate” como melhor “software” para STT entre os analisados. O fato do “Speech to Text Dictate” ter obtido nota máxima (5) para todos os quesitos subjetivos reforça a convicção de ter sido feita a escolha correta.

Apesar da qualidade dos “softwares” para conversão de texto para escrita atuais, eles são pouco utilizados para a melhoria da qualidade de vida dos surdos, esperamos que esse trabalho contribua para efetiva utilização dessa classe de “software”.

Certamente a dificuldade na interpretação da língua escrita é um limitador para a aplicação dessa tecnologia para nascidos surdos, no entanto, em nossa avaliação, os “softwares” para STT podem trazer grandes ganhos para pessoas que tenham ficado surdas após a alfabetização, e pode trazer uma motivação a mais para que surdos de nascença melhorem suas habilidades na utilização do português escrito.

Como trabalho futuro, destacamos a importância da realização de testes em situações reais de utilização, com surdos. Testes dessa natureza são de extrema importância, pois ouvintes têm dificuldades de avaliação das dificuldades dos surdos devido à distância entre suas realidades de vida.

Como podemos observar pela tabela 2 (ver Tabela 2), houve pouca diferença entre os aplicativos avaliados, nos quesitos Instalação, usabilidade e documentação. Tal fato deve-se à predominância da utilização de smartphones como plataforma de execução. As plataformas de distribuição e os ambientes de desenvolvimento para Android e IOS tendem a padronizar as formas de instalação, documentação e utilização desses “softwares”, portanto, os softwares analisados obtiveram avaliações semelhantes.

Tabela 2: avaliação geral.

	Instalação	Usabilidade	Documentação	Tradução
--	------------	-------------	--------------	----------

				WER	CCP
Dictate	5	1	4	0,024	17
HandTalk	5	3	5	0,04	31
Prodeaf	5	5	5	0,024	15
Speech to Text Dictate	5	5	4	0,008	1
Speechy	5	5	4	0,032	7
Google Board	5	5	4	0,016	20
Listnote	5	5	4	0,008	10
Speechnotes:	5	5	4	0,008	5
Speech To Text / Text To Speech	5	2	4	0,008	5
Speech To Text Translator (TTS):	5	5	4	0,008	10

Seguindo este método de avaliação, observamos os aplicativos com maior pontuação total nos itens: Instalação; Usabilidade e Documentação. E com menor pontuação total nos critérios WER e CCP. Cruzando esses dois parâmetros concluímos que o aplicativo “Speechy to Text Dictate” foi eleito o melhor com 14 estrelas, e apenas 1,008 de erro de tradução (WER + CCP). Ele foi seguido pelo “Speechnotes” também com 14 estrelas, mas com 5,008 de erros de tradução. Em seguida temos o “Speechy” com 14 estrelas, e com 7,032 de erro de tradução.

8. Referências

- Alain, L. Y Vejarano, R. (2016). Alternativas tecnológicas para mejorar la comunicación de personas con discapacidad auditiva en la educación superior panameña. *Revista de Educación de la Universidad de Granada*, 23: 219-235.
- Barreto SS, Ortiz KZ (2010). Inteligibilidade: efeitos da análise de transcrição e do estímulo de fala. *Pró-Fono Revista de Atualização Científica*.
- Carneiro, Carlos Felipe Rocha (2016). *Hermes: um modelo para acessibilidade ubíqua dedicado à deficiência auditiva*. Universidade do vale do Rio dos Sinos, UNISINOS. São Leopoldo, Rio Grande do Sul.
- Furlan, Anderson Luís (2016). *Desenvolvimento De Um Protótipo De Aplicativo Móvel Para Conversão De Voz Em Texto E Texto Em Voz, Orientado Ao Apoio À Comunicação De Deficientes Auditivos*. Universidade Federal de Santa Catarina. Aranguá.
- Furui, S. (2007). Avaliação do reconhecimento de fala e falante. Em: Dybkjær, L., Hensen, H., Minker, W., Ide, N. (eds.) *Avaliação de Sistemas de Texto e Fala, Texto, Fala e Tecnologia da Linguagem*, vol. 37, pp. 1–27. Springer, Holanda.

- Grimes, C. H. (2017). Comparative analysis between text translation tools for Brazilian Sign Language. 50 p. Final Project (Bachelor of Science in Computer Science) – State University Northern of Parana, Bandeirantes–PR.
- John, Edward Simon (2016). Interação multimodal adaptativa embarcada em robótica assistiva para comunicação com pssoas com deficiência. Dissertação (mestrado) Univrsidad do Vale dos Sinos- São Leopoldo, Rio Grand do Sul.
- Lasecki, W. S., Miller, C. D., Naim, I., Kushalnagar, R., Sadilek, A., Gildea, D., & Bigham, J. P. (2017). Scribe: deep integration of human and machine intelligence to caption speech in real time. *Communications of the ACM*, 60(9), 93-100
- Martins Santo, Ana Luisa (2017). Speech to Text “software” Evaluation Report.
- Nobuyasu, Itoh; Gakuto, Kurata; Ryuki, Tachibana (2014). Uma Métrica Para Avaliar A Precisão Do Reconhecimento Do Discurso Baseada Na Percepção Do Humano. 研究報告音声言語情報処理 (SLP), v. 2014, n. 11, p. 1-4.
- Pivetta, Ulbricht e Savi. (2011). Tradutores Automáticos Da Linguagem Português Oral E Escrita Para Uma Linguagem Visual-Espacial Da Língua Brasileira De Sinais. Congresso Nacional de hipermídia para aprendizagem, Pelotas RS.
- Santos e Dantas (2017). Tecnologias Assistivas E A Inclusão Do Estudante Surdo Na Educação Superior.Rev. Inter. Educ. Sup. Campinas, SP v.3 n.3 p.494-514 dez.
- S. Stuker, K. Kilgour, and F. Kraft (2012). Quaero 2010 speech-to- text evaluation systems. In High Performance Computing in Science and Engineering.
- Venkateswarlu, R. L. K., Raviteja, R. & Rajeev, R. (2012). The Performance Evaluation of Speech Recognition by Comparative Approach, Advances in Data Mining Knowledge Discovery and Applications.