



Sistema computacional livre para síntese de voz a partir de texto

Fonseca, W. D'A.; **Circe, B.**; **Jacomussi, L.**; Mareze, P. H.



UFSC



Sistema computacional livre para síntese de voz a partir de texto



Fonseca, W. D'A.;
Circe, B.; Jacomussi, L.;
Mareze, P. H.



Realização



SBFa

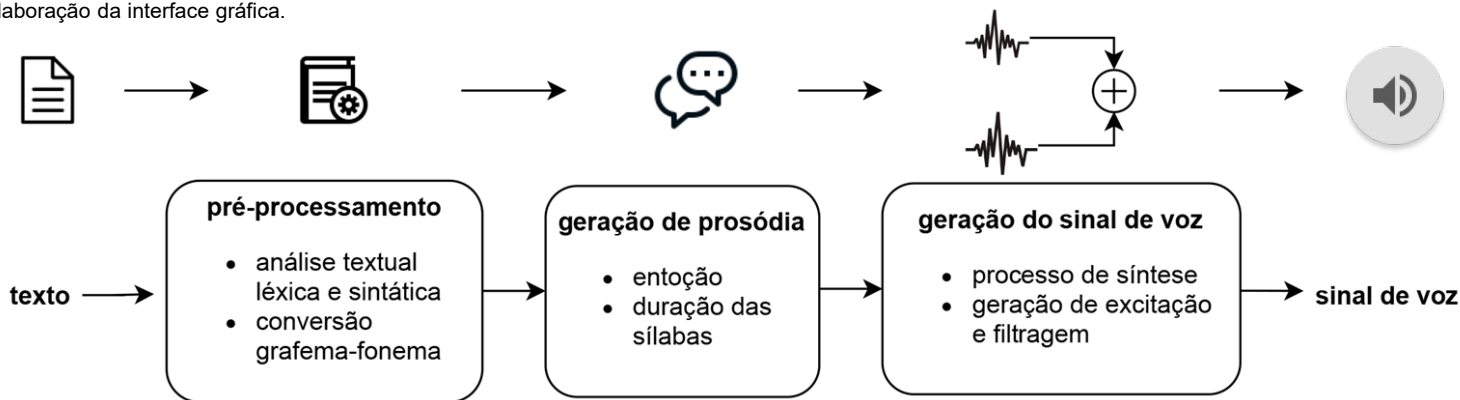
Organização



INTRODUÇÃO: Sistemas de voz têm sido cada vez mais presentes em nossas vidas. Com isso, é importante facilitar o acesso de forma gratuita e amigável à ferramenta de síntese e reprodução de voz a partir de texto.

OBJETIVO: Este trabalho possui o objetivo de apresentar e oferecer facilidades para o uso de um *software* (com bibliotecas livres, *freeware*) para síntese (e reprodução) de voz a partir de texto, auxiliando a comunidade acadêmica em outras pesquisas e/ou no desenvolvimento de outras tecnologias assistivas.

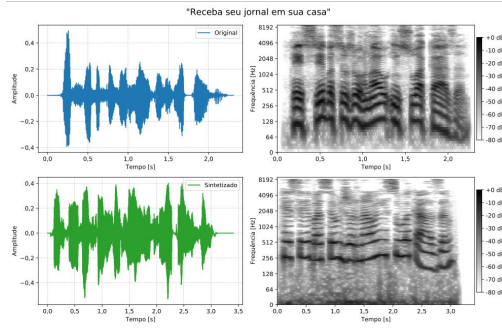
MÉTODO: O *software* é desenvolvido em linguagem de programação Python, sendo as bibliotecas gTTS (Google text-to-speech) e PyQt5 os elementos centrais na constituição dos códigos desenvolvidos. Elas são utilizadas, respectivamente, para construção dos algoritmos de síntese de voz e para a elaboração da interface gráfica.



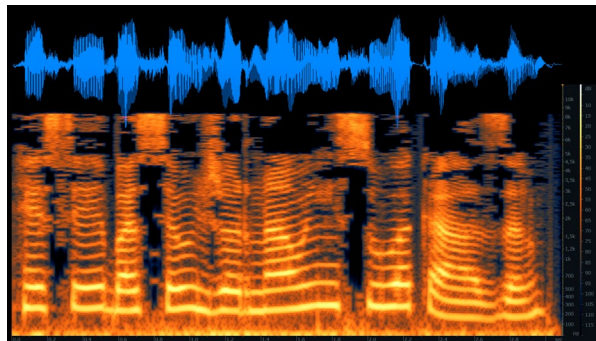
"Receba seu jornal
em sua casa"



Comparação



Sintetizado



Arquivo wav anexo a este PDF.

Sistema computacional livre para síntese de voz a partir de texto



Fonseca, W. D'A.;
Circe, B.; Jacomussi, L.;
Mareze, P. H.



Realização

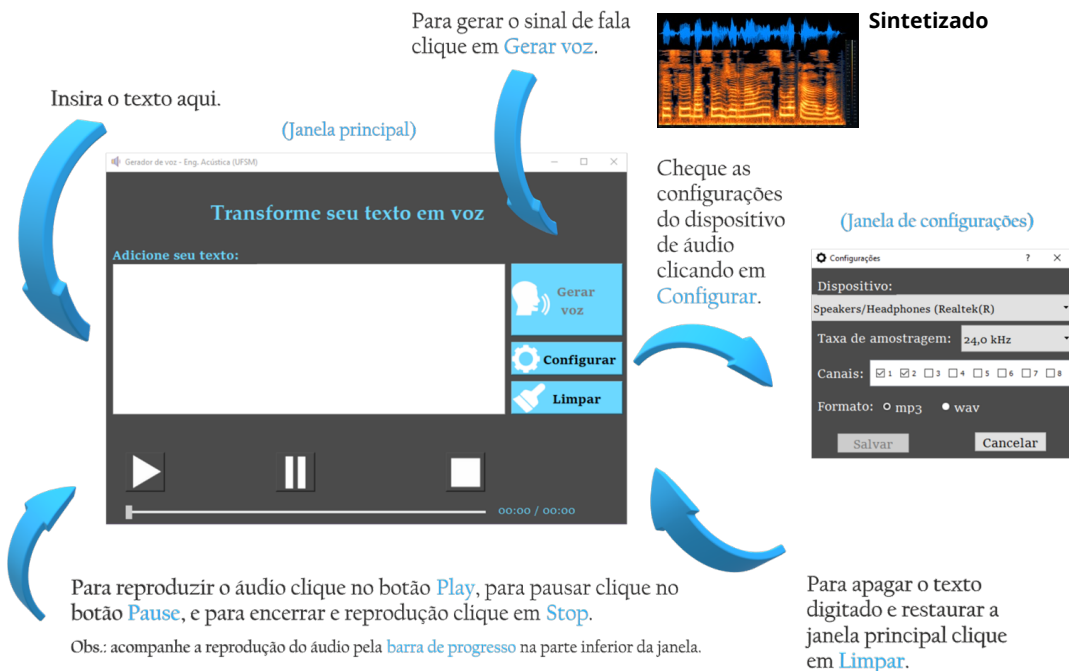


SBFa

Organização



RESULTADOS: A janela principal da interface do aplicativo pode ser vista à esquerda. À direita, vemos a janela de configurações, em que é possível escolher por qual dispositivo o áudio será reproduzido, bem como a sua taxa de amostragem, o número de canais e a opção do formato, podendo ser mp3 ou wav. Para utilizar o aplicativo, o usuário deverá inserir um texto, checar as configurações do dispositivo de áudio no botão “Configurar” e clicar em “Gerar voz”, na janela principal. Por fim, para ouvir o áudio sintetizado, basta clicar no botão “Play”.



É possível ainda acompanhar a reprodução do áudio na barra de progresso na parte inferior da janela principal, podendo pausar pelo botão “Pause” e encerrar a reprodução pelo botão “Stop”. Para deletar o texto inserido e restaurar a janela principal, basta clicar no botão “Limpar”.

O código fonte do aplicativo, bem como informações sobre a instalação e utilização podem ser encontradas em [4].



Disponível gratuitamente em:

<https://github.com/leonardojacomussi/speech-synthesizer>

CONCLUSÃO: O aplicativo apresenta como vantagens sua interface gráfica de fácil uso e sua versatilidade na escolha do dispositivo de áudio, além da possibilidade de configurar parâmetros do arquivo de áudio como taxa de amostragem, número de canais e formato.

Referências:

- [1] Fang Chen (ed.): Speech Technology: Theory and Applications. Springer, 2010.
- [2] Möller, Sebastian. Módulo 6: Automatic Speech Recognition and Text-to-Speech Synthesis. edX courses. Applications in Communication Acoustics, 2020.
- [3] PyQt5 Reference Guide - PyQt v5.15.0 Reference Guide. (Acesso em: 31 de maio de 2020). Disponível em: <https://www.riverbankcomputing.com/static/Docs/PyQt5>
- [4] Speech-synthesizer (Acesso em: 15 de setembro de 2020). Disponível em: <https://github.com/leonardojacomussi/speech-synthesizer>

Sistema computacional livre para síntese de voz a partir de texto

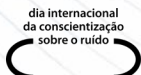


Fonseca, W. D'A.;
Circe, B.; Jacomussi, L.;
Mareze, P. H.



EAC
ENGENHARIA ACÚSTICA | UFSM

Apoio



Realização



Organização



Sistema computacional livre para síntese de voz a partir de texto

INTRODUÇÃO: Sistemas computacionais de síntese de voz têm alcançado maior robustez e sofisticação, especialmente na última década. Como consequência disso, tecnologias da fala sintética estão cada vez mais presentes e acessíveis (em diversos dispositivos) na vida moderna. De forma rudimentar, eles podem ser explicados como sistemas autônomos que são capazes de “ler textos”, convertendo as informações gráficas em sonoras.

OBJETIVO: Este trabalho possui o objetivo apresentar e oferecer facilidades para o uso de um software (com bibliotecas livres, freeware) para síntese (e reprodução) de voz a partir de texto. Para este fim, uma interface gráfica (GUI, graphical user interface) foi desenvolvida, facilitando o uso também por leigos em programação. O software como um todo pode ser aplicado em outras pesquisas e/ou no desenvolvimento de outras tecnologias assistivas.

MÉTODO: O software é desenvolvido em linguagem de programação Python, que é livre para uso (ou seja, sem custo para desenvolvedor e usuário) e independente de sistema operacional (isto é, pode rodar em Windows, MacOS e Linux). Esses são aspectos importantes ressaltados neste projeto. As bibliotecas utilizadas gTTS (Google text-to-speech) e PyQT5 são os elementos centrais na constituição dos códigos desenvolvidos. Elas são utilizadas, respectivamente, para construção dos algoritmos de síntese de voz e para a elaboração da interface gráfica (tornando assim o software mais amigável).

RESULTADOS: A construção e configuração deste sistema computacional livre (freeware) resulta em uma ferramenta gratuita de síntese de fala, que pode ser usada por profissionais e/ou estudantes. Ela está disponibilizada online na plataforma de hospedagem de software GitHub (que detém todas as informações necessárias para uso, além dos códigos comentados). Apesar de se saber que vozes sintéticas têm limitações no que concerne à compreensão do receptor, estima-se que ela pode ajudar em estudos de reabilitação e/ou na assimilação de texto-fala.

CONCLUSÃO: A criação de um sistema de síntese de voz acessível e de uso livre é de considerável importância, sobretudo para a comunidade acadêmica (e/ou aqueles que tem restrições de capital). Os áudios gerados podem ser tanto utilizados de forma praticamente instantânea, quando gravados em um arquivo de áudio como wave ou mp3, por exemplo. Como objetivo futuro, espera-se expandir o projeto, adicionando mais recursos que possam ser úteis nas práticas de fonoaudiologia e engenharia.

Referências:

- [1] Fang Chen (ed.): Speech Technology: Theory and Applications. Springer, 2010.
- [2] Möller, Sebastian. Módulo 6: Automatic Speech Recognition and Text-to-Speech Synthesis. edX courses. Applications in Communication Acoustics, 2020.
- [3] PyQt5 Reference Guide - PyQt v5.15.0 Reference Guide. (Acesso em: 31 de maio de 2020). Disponível em: <https://www.riverbankcomputing.com/static/Docs/PyQt5>
- [4] Speech-synthesizer (Acesso em: 15 de setembro de 2020). Disponível em: <https://github.com/leonardojacomussi/speech-synthesizer>



Disponível gratuitamente em:

<https://github.com/leonardojacomussi/speech-synthesizer>