

Tecnologias assistivas para surdos

1st Hiel Alves Rocha

O Instituto de Educação Superior de Brasília (IESB)
Ciência da Computação (CCO)
Brasília, Brasil
hiel.rocha@iesb.edu.br

2nd Iago Alves Abadia dos Santos

O Instituto de Educação Superior de Brasília (IESB)
Ciência da Computação (CCO)
Brasília, Brasil
iago.santos@iesb.edu.br

3rd Matheus Hipólito da Mata

O Instituto de Educação Superior de Brasília (IESB)
Ciência da Computação (CCO)
Brasília, Brasil
matheus.mata@iesb.edu.br

4th Gustavo Alves de Souza Marques

O Instituto de Educação Superior de Brasília (IESB)
Ciência da Computação (CCO)
Brasília, Brasil
gustavoalvessm21@gmail.com

Abstract—Inclusion is a topic of great importance that has been gaining prominence over time. When it comes to accessibility, there is still a long way to go and changes that must be made in the long term. In colleges and universities, the number of people with disabilities grows, among them, people with some degree of deafness. In this context, there are numerous situations in which the deaf person finds access difficult due to little or no measure that promotes accessibility. Seminars, conferences and similar events do not always have Libras interpreters and, depending on the size of the venue, the distance between the interpreter and the deaf person can be too great for the translation to be properly understood. Bearing this in mind, this article proposes the use of the "Applications" platform as an alternative to provide an interpreter or to complement the means of access to the content of these events for people who can see the translation in Libras or who, eventually, have Mastery of Sign Language.

I. INTRODUÇÃO

O uso das tecnologias de informação e comunicação (TICs) acessíveis às pessoas surdas é um tema que vem crescendo no Brasil, principalmente desde 2000, com o reconhecimento da Língua Brasileira de Sinais (Libras), por meio da Lei 10.436/2002 (BRASIL, 2002) e a criação da Lei de Inclusão de Pessoas com Deficiência (PCD) – 13.146/2015, que aborda, entre outros aspectos, o direito à informação, à cultura e ao lazer em formato acessível. Essas regulamentações chamaram a atenção para as peculiaridades linguísticas dos surdos, que, segundo o último IBGE.1 censo demográfico, representam 7,6A Lei 10.436/2002, por exemplo, estipula que o poder público - incluindo as concessionárias de serviços públicos - deve apoiar o uso e a divulgação da Libras como ferramenta de comunicação objetiva e de uso corrente (BRASIL, 2002). O Decreto 5.626, de 22 de dezembro de 2005, que veio a regulamentar essa normativa, determinou a inclusão da Libras como disciplina obrigatória nos cursos de Fonoaudiologia e Formação de Professores para o Ensino Médio e Superior; e como disciplina optativa nos demais cursos, em todas as instituições de ensino do país, públicas e privadas. As concessionárias de serviços públicos e toda a administração pública também tiveram que prestar atendimento em Libras com pelo

menos 5Da mesma forma, em 2014, a Agência Nacional do Cinema (Ancine), por meio de instrução normativa nº. 116, instituiu obrigatoriedade de legendagem, audiodescrição e Libras em todos os projetos de produção audiovisual financiados com recursos públicos federais geridos pela Agência, a fim de promover a acessibilidade (BRASIL, 2014). Embora, do ponto de vista clínico ou orgânico, a surdez seja definida como a redução da capacidade de perceber sons (RINALDI et al, 1997), também pode ser entendida, além da deficiência auditiva, como uma marca identitária de um grupo cultural específico. Nesse entendimento, os surdos são "pessoas que não se consideram deficientes, utilizam a língua de sinais, valorizam sua história, arte e literatura e propõem uma pedagogia adequada para a educação das crianças surdas" (BISOL; VALENTINI, 2011). De acordo com Sacos (1998), a linguagem dos sinais é a forma natural de comunicação desses sujeitos, uma vez que sua modalidade gesto-visual a torna acessível em comparação com as línguas orais, como o português. Criada na segunda metade do século 19, a Libras só foi considerada oficial em 2002, quando os surdos brasileiros passaram a ter duas línguas oficiais, sendo a primeira a libra e a segunda, o português (BRASIL, 2002; MONTEIRO, 2006). As dificuldades de compreensão dessa segunda língua, devido às suas diferenças estruturais e significativas em relação à Libras, constituem uma barreira de acesso que muitas vezes não é percebida pelos produtores de sites e conteúdo digital, como observado por Vai e Gomes (2011). Segundo esses autores, quando examinamos os recursos de acessibilidade disponíveis para surdos na rede mundial de computadores, geralmente temos apenas vídeos com legendas explicativas em português – o que não contempla a realidade desse grupo social, que utiliza o português como segunda língua. Ressaltam que se deve ao senso comum de que as pessoas surdas não ouvem, mas podem ler e compreender textos escritos em português porque a leitura visual estaria relacionada a um sentido que funciona para elas; e sugerir o uso de tradutores e vitrines Libras, além da produção de conteúdo nessa língua, para que as pessoas surdas possam ter acesso a uma maior diversidade

de produtos de comunicação para consumir, contribuindo de forma mais efetiva para a superação da barreira linguística para sua efetiva inclusão na sociedade. Visando essa demanda, diversas TICs surgiram com a proposta de minimizar as barreiras de comunicação em diferentes mídias, oferecendo acesso mais amplo para os surdos, por exemplo, ao cinema, à televisão e à internet. Em 2013, surgiu a primeira emissora de TV para surdos, a TV INES, fruto de uma parceria entre o Instituto Nacional de Educação de Surdos (INES) e a Associação Roquette Pinto de Comunicação Educativa (ACERP). Com uma equipe de seis apresentadores surdos e cinco intérpretes de Libras, oferece programação de filmes, notícias, desenhos animados, esportes, cultura e tecnologia, que é distribuída via satélite para antenas parabólicas, TVs a cabo, smartphones e tablets. Tudo é feito em Libras e traduzido para o português com legendas (SIQUEIRA; SOUZA, 2016). Alguns estudos, como Amorim et al (2010), apontam possibilidades de tradução das legendas ocultas para Libras, criando sinais na extensão animada GIF. Além disso, existem tradutores que podem ser usados em dispositivos móveis e totens, como o internacionalmente premiado ProDeaf2e HandTalk, considerado pela Organização das Nações Unidas em 2013 como o melhor aplicativo de inclusão social do mundo³. Inserir essas tecnologias na produção de conteúdo dos mais diversos gêneros é um desafio no campo da Comunicação, pois é preciso estar preparado não apenas para superar barreiras linguísticas, mas também para atender às demandas do público surdo por informações de alta qualidade e conectadas com suas culturas e interesses, permitindo que os surdos pertençam a um grupo específico onde possam iniciar a construção de fronteiras identitárias com o outro, bem como obter o reconhecimento social dos demais grupos, [uma vez que] a constituição da identidade pela pessoa surda está relacionada à presença de uma língua que lhe dá a possibilidade de se constituir como um "sinal" que pode estabelecer práticas discursivas e sociais

(QUADROS; STUMPF, 2009, p.171).

II. METODOLOGIA

Esta seção ilustra o quadro do sistema de aprendizagem de língua gestual proposto para a LSB (Língua de Sinais Brasileira). O modelo proposto é subdividido em três módulos, conforme ilustrado. O primeiro módulo corresponde à conversão da fala para uma sentença em português, que é então processada usando PNL para obter a sentença LSB correspondente. Por fim, alimentamos a sentença LSB extraída para o modelo de avatar para produzir a respectiva língua de sinais.

Tradução da Sentença LSB da Sentença em português Esta seção fornece os detalhes do processo de conversão do texto em português para o texto LSB correspondente. As palavras da sentença LSB foram identificadas para gerar movimentos de sinais correspondentes. Para a conversão do português para o LSB, utilizamos o Java. O modelo de conversão da sentença LSB de uma sentença em português é plotado.

Pré-processamento de texto de entrada usando expressão regular Se um usuário inserir por engano uma palavra inválida/incorreta, a função "editar distância" será usada para

obter uma palavra válida equivalente. Alguns exemplos de palavras com erros ortográficos, juntamente com as palavras válidas correspondentes.

A função de distância de edição usa duas cadeias de caracteres (origem e destino) e modifica a cadeia de caracteres de origem de forma que as cadeias de caracteres de origem e de destino se tornem iguais. O Java divide a frase em português em pares de palavras separadas – PoS usando o tokenizador de texto. A expressão regular identifica a frase significativa em português usando a regra do léxico. Durante o pré-processamento do texto de entrada, definimos a expressão regular (1) usando os tokens PoS do módulo Java. A expressão regular começa com pelo menos uma frase verbal (VP) e termina com uma frase nominal (NP). Na parte intermediária, a expressão regular pode ter zero ou mais número de quaisquer palavras que correspondam a tokens PoS (preposição (PP) ou um pronome (PRP) ou adjetivo (JJ)). Em uma expressão regular, + refere-se a um ou mais símbolos, enquanto refere-se a zero ou mais símbolos. Portanto (VP)+

representa uma ou mais frases verbais. Por exemplo, nossa primeira frase, "Venha para minha casa", começa com a frase verbal ("vir"), seguida por uma preposição ("jogar"), pronome ("você"), e termina com uma frase nominal ("futebol").

(VP)+(PP—PRP—JJ)(NP) (1) onde VP (VB,VBN), NP (NN), VB ("Olá","Obrigado","Por favor"), VBN ("venha"), PP ("para","com"), PRP ("meu","você","me"), JJ ("Bom"), NN ("casa", "manhã").

Análise Sintática e Estrutura Lógica Após a etapa de pré-processamento, o módulo java retorna a árvore de análise com base nos tokens gramaticais (VP, PP, NP, etc.). Em seguida, construímos a árvore de derivação da Gramática Livre de Contexto (CFG), que é semelhante à árvore de análise do módulo java. CFG consiste em símbolos variáveis/não terminais, símbolos terminais e um conjunto de regras de produção. Os símbolos não terminais geralmente aparecem no lado esquerdo das regras de produção, embora também possam ser introduzidos no lado direito. Os símbolos terminais aparecem no lado direito das regras. A regra de produção gera a cadeia de caracteres do terminal a partir do símbolo não terminal. A árvore de derivação pode representar a derivação da cadeia de caracteres terminal a partir do símbolo não-terminal. Na árvore de derivação, símbolos terminais e não terminais referem-se às folhas e nós intermediários da árvore. Cada sentença LSB significativa tem sua própria árvore de derivação. Após a criação da árvore de derivação, as folhas da árvore são combinadas para fazer uma estrutura lógica para a língua de sinais.

Gerador de Scripts e Sentença LSB O gerador de scripts cria um script para gerar uma sentença LSB a partir da sentença em português. O script usa uma frase válida em português (após análise semântica) como entrada e gera a árvore de sequência, onde cada nó da árvore é relacionado a diferentes gestos que estão associados ao movimento do avatar. A árvore de sequência mantém a ordem do movimento realizado pelo modelo de avatar. As estruturas das sentenças português e LSB são bem diferentes. A representação da LSB

a partir das sentenças em português é feita usando a Gramática Funcional Lexical (LFG). A estrutura f do LFG codifica a relação gramatical, como sujeito, objeto e tempo de uma frase de entrada. A LSB segue a ordem das palavras "Sujeito-Objeto-Verbo", enquanto a língua inglesa segue a ordem das palavras "Sujeito-Verbo-Objeto". Além disso, a sentença LSB não considera qualquer conjunção e preposição na sentença.

Geração do Movimento de Sinais A geração de movimentos de signos com base no texto de entrada é realizada com a ajuda de uma ferramenta de animação chamada Blender. A ferramenta é popularmente usada para projetar jogos, animação 3D, etc. A lógica do jogo e o objeto do jogo são os principais componentes do motor de jogo do Blender. Desenvolvemos o avatar 3D determinando sua forma geométrica. Todo o processo de criação do avatar é dividido em três etapas. Primeiro, o esqueleto e o rosto do avatar são criados. Na segunda etapa, definimos o ponto de vista ou orientação do modelo. Na terceira etapa, definimos as articulações de movimento e expressões faciais do avatar. Em seguida, fornecemos a sequência de quadros que determinam o movimento do avatar sobre a determinada sequência de palavras ao longo do tempo. Finalmente, o movimento (movimento como andar, mostrar figuras, mover as mãos, etc.) é definido dando uma animação sólida. O motor de jogo foi escrito a partir do zero em C++ como uma parte principalmente independente e inclui suporte para recursos como script Java e som OpenAI 3D. Neste terceiro módulo, geramos movimentos de sinal para a sentença Libras (gerada no segundo módulo). Toda a estrutura da geração do movimento do avatar a partir da sentença LSB. Para a geração do movimento do sinal a partir da LSB, inicialmente, os parâmetros de animação são extraídos da sentença LSB. Uma vez que os parâmetros de animação são identificados, a lista de movimento de cada signo é realizada usando um modelo de avatar 3D. No modelo de avatar 3D proposto, cada movimento é associado a vários movimentos, e todos esses movimentos são listados no arquivo "motionlist". Uma variável de contador é inicializada para controlar o movimento atual. Cada movimento tem seus carimbos de data/hora mencionando por quanto tempo a ação de diferentes gestos será executada. Cada movimento é gerado por um atuador de movimento específico quando algum evento do sensor ocorre. O controlador atua como um ligador entre o sensor e o atuador. O loop condicional verifica os limites máximos do número de movimentos, e ele executa o movimento um a um usando um atuador específico (por exemplo, o atuador [contador]). O próximo movimento válido é executado incrementando a variável do contador. Se o valor da variável contador exceder o número de movimentos no arquivo "motionlist", a variável contador junto com a "motionlist" será redefinida para o valor padrão e o próximo movimento será executado.

III. RESULTADO

Aqui, apresentamos uma análise detalhada do sistema de aprendizagem de língua de sinais proposto usando o modelo 3D Avatar. Esta seção é composta por quatro seções, a saber: banco de dados de sinais (1), resultados do processo de

tradução (2) e geração do movimento de sinais a partir da sentença LSB (3).

(1) **Banco de Dados de Sinais** Criamos um banco de dados de linguagem de sinais que contém frases baseadas em 50 palavras LSB usadas diariamente (por exemplo, eu, meu, vem, casa, bem-vindo, desculpe, chuva, você, bebê, vento, homem, mulher, etc.) e outros diálogos entre diferentes usuários. Criamos 150 frases que contêm 763 palavras diferentes, incluindo as palavras mais usadas na LSB. Para cada palavra, os movimentos dos sinais foram definidos no kit liquidificador. A descrição do banco de dados de sinais está representada na Tabela 4. Os itens de vocabulário foram criados com base nas palavras únicas na LSB. Para melhor compreensão, representamos quatro sequências animadas de cada palavra. Para simplificar, apresentamos alguns exemplos de sequências de movimento de sinais usando duas séries animadas. A tabela mostra os movimentos dos sinais juntamente com suas palavras reais em inglês. Todos esses movimentos de sinais foram definidos com a ajuda de um especialista em língua de sinais de uma escola para deficientes auditivos.

(2) **Resultados do Processo de Tradução** Esta seção ilustra o processo de tradução do modelo proposto. O processo de tradução inclui conversão de sentenças do inglês para LSB e representação de sentença para assinatura LSB. A avaliação do sistema proposto foi realizada dividindo-se as sentenças geradas em uma proporção de 80:20 entre as séries de treinamento e teste, respectivamente, e a Taxa de Erro de Palavra (WER) da palavra digitada foi registrada. O resultado do sistema de processamento de texto é apresentado na Tabela 7, onde a métrica WER é derivada da distância de Levenshtein (função editar distância). Aqui comparamos a palavra da frase de referência e a sentença de saída. A distância calcula o número de edições/alterações (inserção/exclusão e modificações) necessárias para converter o texto de entrada para o texto de referência correto. Na Tabela 7, Ins, Del, e Sub referem-se ao número de operações de inserção, exclusão e modificação/substituição para converter o texto de origem para o texto de destino adequado, respectivamente.

(3) **Geração do Movimento de Sinais a partir da Sentença LSB** Após a conversão da sentença LSB da sentença em português, procedeu-se à geração do movimento do sinal para a sentença LSB. O avatar gera movimentos de signos animados para cada palavra significativa. Aqui, usamos o avatar usando o software blender. Traçamos todos os movimentos da LSB correspondentes às frases em Português. A descreve as ações de representação em língua de sinais para a frase em português "você". representa a representação em língua de sinais da frase em português "olhar", e a descreve a representação em língua de sinais da frase em português "futebol". representa a representação em língua gestual da frase portuguesa "Você olhar futebol".

CONCLUSÃO

Neste trabalho, desenvolvemos um sistema de aprendizagem de linguagem de sinais baseado em avatar 3D que converte a texto em em movimentos LSB correspondentes.

Inicialmente, a fala de entrada é convertida em uma frase equivalente em Português usando o serviço IBM-Watson. O texto em Português convertido é posteriormente traduzido para a sentença LSB correspondente usando expressão regular e gerador de script. Finalmente, cada palavra da sentença LSB é transformada em um movimento de sinal equivalente, representado por um avatar 3D. O método de tradução de LSB foi avaliado pelas métricas SER, BLEU e NIST. O modelo proposto obteve um escore BLEU de 82,30, que representa a precisão da tradução de sentenças do Português para o LSB. O modelo de tradução texto-sinal (da sentença LSB para o movimento do sinal) obteve um escore de 10,50 RSE, o que significa que 89,50 dos sinais foram gerados corretamente pelo modelo avatar 3D para a respectiva sentença LSB. Pode-se notar que o sistema proposto foi desenvolvido para um corpus limitado, e nenhuma expressão facial foi incluída, o que pode ser considerado como uma parte importante de qualquer sistema de língua de sinais. A transição entre os signos durante a execução de uma frase de sinal pode ser melhorada ainda mais aprendendo transições específicas com base nas posições das mãos durante a assinatura. Além disso, o sistema de reconhecimento da língua de sinais que converte um sinal em texto é significativamente mais difícil de desenvolver. Tal sistema pode ser adicionado dentro da estrutura proposta para construir um sistema completo de interpretação da língua de sinais

REFERENCES

- [1] <https://www.autodesk.com.br/products/3ds-max/overview?term=1-YEARtab=subscription>
- [2] <https://www.opengl.org/>
- [3] <https://www.libras.com.br/signwriting>
- [4] BRASIL. Decreto nº 5.626, de 22 de dezembro de 2005 Regulamenta a Lei no 10.436, de 24 de abril de 2002. <http://www.planalto.gov.br/ccivil03/a/to2004-2006/2005/decreto/d5626.htm>
- [5] Lei nº 10.436, de 24 de abril de 2002 Dispõe sobre a Língua Brasileira de Sinais – Libras e dá outras providências. <http://www.planalto.gov.br/ccivil03/leis/2002/L10436.htm>
- [6] Lei nº 13.146, de 03 de julho de 2015 Institui a Lei brasileira de inclusão da pessoa com deficiência (Estatuto da Pessoa com Deficiência). <http://www.planalto.gov.br/ccivil03/a/to2015-2018/2015/lei/l13146.htm>
- [7] MONTARDO, S. P.; PASSERINO, L. M. Inclusão social via acessibilidade digital: proposta de inclusão digital para pessoas com necessidades especiais. Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação Belo Horizonte, v.8, 2007. <http://compos.org.br/seer/index.php/e-compos/article/view/144/145>
- [8] QUEIROZ, L. S.; RÚBIO, J. A. S. A aquisição da linguagem e a interação social: a LIBRAS como formadora da identidade do surdo. Revista Eletrônica Saberes da Educação, São Roque, v.5, n.1. 2014. http://www.facsao Roque.br/novo/publicacoes/publicacao_tual_2014.html
- [9] RAMOS, C. R. LIBRAS: A língua de Sinais dos Surdos brasileiros, 2006. <http://www.editora-arara-azul.com.br/pdf/artigo2.pdf>
- [10] SIQUEIRA, J. M.; SOUZA, J. B. de. Jornalismo e acessibilidade: TV INES, primeira webtv acessível do Brasil. In: CONGRESSO DAS CIÊNCIAS DA COMUNICAÇÃO NA REGIÃO NORDESTE. Caruaru, 7 a 9 jul. 2016. Anais... <http://www.portalintercom.org.br/anais/nordeste2016/resumos/R52-1460-1.pdf>
- [11] VENTURA, L. A. S. Cartilha de acessibilidade na web Blog Vencer Limites/O Estado de S. Paulo, 14/04/2015. <http://brasil.estadao.com.br/blogs/vencer-limites/cartilha-de-acessibilidade-na-web/>
- [12] WORLD WIDE WEB Consortium Brasil (WC3). Cartilha Acessibilidade na WEB, 2015. <http://www.w3c.br/pub/Materiais/Publicacoes/W3C/cartilha-w3cbr-acessibilidade-web-fasciculo-I.html>