

## ENTRE GESTOS E ALGORITMOS: INTELIGÊNCIA ARTIFICIAL E AS NOVAS FRONTEIRAS DA FONOLOGIA VISUAL

### BETWEEN GESTURES AND ALGORITHMS: ARTIFICIAL INTELLIGENCE AND THE NEW FRONTIERS OF VISUAL PHONOLOGY

Denis Ramón Funes Flores <sup>1</sup>

Universidade Federal do Tocantins

**Resumo:** Este artigo investiga as interfaces entre a fonologia visual das línguas de sinais e os sistemas algorítmicos baseados em inteligência artificial (IA), com foco na Língua Brasileira de Sinais (Libras). Parte-se do pressuposto de que, embora a IA tenha avançado significativamente no reconhecimento de gestos, muitos modelos ainda negligenciam os parâmetros fonológicos fundamentais — como configuração de mão, movimento, localização, orientação e expressões não manuais — que constituem a estrutura linguística das línguas de sinais. Por meio de uma revisão teórico-crítica, analisam-se estudos recentes em linguística de línguas de sinais e em processamento multimodal, buscando identificar lacunas conceituais e técnicas na integração entre teoria linguística e modelagem computacional. Os resultados indicam que a maioria dos sistemas de IA opera com uma visão instrumentalizada do gesto, desconsiderando seu estatuto linguístico e semiótico, o que pode reforçar desigualdades no acesso à comunicação e à informação para pessoas surdas. Propõe-se, então, a necessidade urgente de corpora linguísticos sinalizados com anotação fonológica rigorosa e de parcerias interdisciplinares que incluam pesquisadores surdos e especialistas em linguística. A reflexão aqui desenvolvida contribui para repensar os fundamentos da “fonologia computacional” em contextos visuais e para orientar o desenvolvimento de tecnologias mais éticas, inclusivas e linguisticamente informadas.

**Palavras-chave:** Fonologia Visual; Línguas de Sinais; Inteligência Artificial; Libras; Análise Multimodal.

**Abstract:** This article explores the interfaces between the visual phonology of sign languages and algorithmic systems based on artificial intelligence (AI), with a focus on Brazilian Sign Language (Libras). While AI has made significant advances in gesture recognition, many models still overlook core phonological parameters—such as handshape, movement, location, orientation, and non-manual markers—that constitute the linguistic structure of sign languages. Through a critical-theoretical review, recent studies in sign language linguistics and multimodal processing are examined to identify conceptual and technical gaps between linguistic theory and computational modeling. Findings suggest that most AI systems operate with an instrumentalized view of gesture, disregarding its linguistic and semiotic status, potentially reinforcing communication and information access inequalities for Deaf individuals. The article argues for the urgent development of linguistically annotated signed corpora and interdisciplinary collaborations that include Deaf researchers and linguists. This reflection contributes to rethinking the foundations of “computational phonology” in visual contexts and to guiding the design of more ethical, inclusive, and linguistically grounded technologies.

**Keywords:** 3 to 5, separated by colon.

---

<sup>1</sup> Pesquisador da Universidade Federal do Tocantins - UFT. Email: derafuflo@mail.ufr.edu.br

**Submetido em 29 de dezembro de 2025.**

**Aprovado em 14 de abril de 2026.**

## **Introdução**

As línguas de sinais, longe de serem meros conjuntos de gestos ou mímicas, constituem sistemas linguísticos plenos, complexos e estruturados, com gramáticas próprias, capazes de expressar abstrações, narrativas e conceitos científicos com a mesma riqueza das línguas orais (STOKOE, 1960; QUADROS; KARNOPP, 2004). No Brasil, a Língua Brasileira de Sinais (Libras) foi reconhecida legalmente como meio de comunicação e expressão da comunidade surda pela Lei nº 10.436, de 2002, consolidando seu estatuto linguístico e cultural. Desde então, a produção acadêmica sobre Libras tem crescido exponencialmente, com destaque para estudos fonológicos que desconstróem visões medicalizantes e patologizantes da surdez, reafirmando a autonomia linguística e a identidade sociocultural dos surdos (PERLIN; QUADROS, 2008; FELIPE, 2012).

Nesse contexto, a fonologia visual emerge como um dos pilares teórico-analíticos centrais na descrição linguística das línguas de sinais. Diferentemente das línguas orais — cujos traços distintivos operam no canal auditivo —, as línguas de sinais organizam-se no espaço visual-gestual, com base em parâmetros articulatórios como configuração da mão, localização, movimento, orientação da palma e expressões não manuais (BRENTARI, 1998; SANDLER; LILLO-MARTIN, 2006). Esses elementos, longe de serem arbitrários, combinam-se de forma sistemática e contrastiva, permitindo a distinção lexical e gramatical — função que, nas línguas orais, é desempenhada pelos fonemas. A compreensão desses mecanismos é, portanto, essencial não apenas para a linguística teórica, mas também para aplicações práticas em educação, tradução e, mais recentemente, em tecnologias digitais.

Nas últimas décadas, o avanço exponencial da inteligência artificial (IA) — especialmente em visão computacional, redes neurais profundas e processamento multimodal — tem impulsionado iniciativas voltadas ao reconhecimento automático de línguas de sinais. Projetos acadêmicos e corporativos vêm desenvolvendo sistemas capazes de “ler” sinais em vídeos, com promessas de tradução em tempo real, interfaces de acessibilidade e assistentes virtuais inclusivos (YIN et al., 2023; FERREIRA et al., 2022). Contudo, muitos desses modelos operam com uma visão reducionista do sinal, tratando-o como um gesto isolado ou uma sequência de poses corporais, sem considerar

sua estrutura fonológica subjacente. Essa desconexão entre modelagem algorítmica e teoria linguística pode comprometer não apenas a precisão técnica, mas também a própria representação da língua sinalizada, correndo o risco de reproduzir vieses algorítmicos e epistemologias coloniais que desconsideram o sujeito surdo como agente linguístico (MIRANDA-GÁLVEZ; ALMEIDA, 2024).

Diante desse cenário, este artigo se propõe a investigar a seguinte pergunta de pesquisa: como a fonologia das línguas de sinais é (ou não) considerada nos modelos algorítmicos atuais baseados em inteligência artificial? O objetivo é realizar uma análise crítica das interfaces entre a fonologia visual e os sistemas computacionais emergentes, identificando lacunas conceituais, metodológicas e éticas na integração entre linguística de línguas de sinais e ciência da computação. A relevância teórica reside na necessidade de consolidar uma fonologia computacional visual que dialogue com os avanços da semântica, pragmática e sociolinguística das línguas de sinais, enquanto a relevância social se expressa no potencial de contribuir para o desenvolvimento de tecnologias verdadeiramente inclusivas, linguisticamente fundamentadas e co-projetadas com a comunidade surda.

O artigo está organizado da seguinte forma: na seção 1, apresentam-se os fundamentos teóricos da fonologia nas línguas de sinais; na seção 2, discute-se o estado da arte da IA aplicada ao reconhecimento de sinais; na seção 3, analisam-se as interfaces e tensões entre fonologia visual e algoritmos; e, por fim, na seção 4, refletem-se as implicações éticas, sociais e linguísticas dessas interações. As considerações finais sintetizam as contribuições e apontam caminhos para pesquisas futuras.

## **1. Fundamentos da fonologia nas línguas de sinais**

A fonologia das línguas de sinais representa um dos pilares teóricos que consolidaram o reconhecimento científico dessas línguas como sistemas linguísticos plenos, desvinculados de quaisquer dependências com as línguas orais. Foi com o trabalho pioneiro de Stokoe (1960) que se estabeleceu a primeira evidência empírica de que a American Sign Language (ASL) possuía uma estrutura fonológica análoga, embora distinta, à das línguas faladas. Ao propor que os sinais poderiam ser decompostos em unidades contrastivas — que ele denominou *cheremes* —, Stokoe inaugurou uma nova era na linguística, demonstrando que a linguagem humana não está

intrinsecamente atrelada ao canal vocal-auditivo, mas sim à capacidade cognitiva de organizar formas simbólicas em sistemas estruturados (STOKOE, 1960).

Com o avanço dos estudos linguísticos, a proposta inicial de Stokoe foi refinada e expandida, especialmente por meio do modelo dos parâmetros fonológicos, amplamente adotado na análise das línguas de sinais. Segundo Brentari (1998), os sinais são compostos por um conjunto de traços articulatorios simultâneos que operam nos domínios manual e não manual, incluindo: (1) configuração da mão (handshape), (2) localização (location), (3) movimento (movement), (4) orientação da palma (orientation) e (5) expressões não manuais (non-manual markers), como sobranceiras, olhar e postura corporal. Esses parâmetros funcionam de maneira combinatória e contrastiva, permitindo distinguir, por exemplo, os sinais de LIBRAS para “mãe” e “pai”, que diferem apenas pela localização do sinal na face (QUADROS; KARNOPP, 2004, p. 56).

Essa estrutura fonológica, no entanto, opera em um canal visual-espacial, o que implica uma organização temporal e espacial distinta da fonologia das línguas orais. Enquanto as línguas faladas articulam fonemas sequencialmente no tempo, as línguas de sinais articulam parâmetros fonológicos de forma simultânea e espacializada, utilizando o corpo e o espaço próximo como suporte articulatorio (SANDLER; LILLO-MARTIN, 2006). Esse aspecto fundamental levou à consolidação do conceito de fonologia visual, termo que ressalta as particularidades modais dessas línguas sem subordiná-las a modelos fonológicos orais (PERLIN; QUADROS, 2008). A fonologia visual, portanto, não é uma mera “adaptação” da fonologia oral, mas um sistema autônomo, com princípios organizacionais próprios, ancorado em propriedades perceptivo-motoras do canal visual-gestual.

No contexto brasileiro, os estudos sobre a fonologia da Libras têm se intensificado nas últimas décadas, especialmente após sua oficialização em 2002. Quadros e Karnopp (2004) ofereceram uma das primeiras descrições sistemáticas dos parâmetros fonológicos da Libras, baseando-se em dados empíricos coletados com surdos sinalizantes nativos. Mais recentemente, pesquisas têm aprofundado aspectos específicos dessa estrutura: Almeida e Souza (2019) analisaram variações fonológicas regionais em sinais de diferentes estados do Brasil; Oliveira e Góes (2020) investigaram o papel das expressões não manuais na marcação de estruturas prosódicas e focais; e Silva et al. (2022) demonstraram, por meio de experimentos psicolinguísticos, como

alterações mínimas em parâmetros fonológicos podem impactar significativamente o reconhecimento lexical por usuários fluentes de Libras.

Esses avanços teóricos e empíricos reforçam uma conclusão central: a fonologia visual não é um detalhe periférico, mas o núcleo estruturante da gramática das línguas de sinais. Ignorar seus princípios em sistemas automatizados — como os baseados em inteligência artificial — é equivalente a tentar analisar uma língua oral sem considerar fonemas, vogais ou consoantes. Para que modelos computacionais alcancem fidedignidade linguística e utilidade prática, é indispensável que incorporem, desde sua concepção, os parâmetros fonológicos como unidades de análise fundamentais, respeitando a simultaneidade, a iconicidade estrutural e a gramaticalização do espaço visual que caracterizam essas línguas (MIRANDA-GÁLVEZ; ALMEIDA, 2024). A próxima seção examinará até que ponto essa incorporação tem sido realizada — ou negligenciada — nas abordagens atuais de IA voltadas ao reconhecimento de línguas de sinais.

## **2. Inteligência artificial e o reconhecimento de línguas de sinais: estado da arte**

Nos últimos anos, o reconhecimento automático de línguas de sinais tem emergido como uma das frentes mais ativas na interseção entre linguística computacional, visão computacional e acessibilidade digital. Impulsionado por avanços em aprendizado profundo (deep learning), estimativa de pose humana (human pose estimation) e modelagem multimodal, o campo tem produzido uma série de abordagens algorítmicas com o objetivo de traduzir sinais para texto ou fala, ou de interpretar sequências sinalizadas em tempo real (YIN et al., 2023; DOGAN et al., 2022). Entre as arquiteturas mais utilizadas destacam-se as redes neurais convolucionais (CNNs), empregadas para extração de características espaciais em quadros de vídeo; as redes recorrentes (RNNs) e transformers, utilizadas para modelar a dinâmica temporal dos sinais; e os modelos baseados em grafos (graph neural networks), que representam o corpo como um conjunto de nós articulados (juntas) e arestas (segmentos corporais), permitindo capturar relações espaciais entre membros (FERREIRA et al., 2022).

Apesar do progresso técnico, uma análise crítica do estado da arte revela uma desconexão persistente entre os objetivos tecnológicos e os fundamentos linguísticos das línguas de sinais. A grande maioria dos sistemas opera sob um paradigma gestualista, no qual o sinal é tratado como um gesto instrumental ou simbólico isolado,

sem considerar sua inserção em uma estrutura gramatical mais ampla (MIRANDA-GÁLVEZ; ALMEIDA, 2024). Por exemplo, muitos modelos de reconhecimento treinam-se para classificar sinais individuais em dicionários estáticos — frequentemente limitados a centenas ou milhares de itens lexicais —, ignorando fenômenos linguísticos essenciais como inflection, agreement, estrutura de tópico-comentário ou o uso de espaço gramatical para referência anafórica (HUENERFORTH et al., 2021). Esse enfoque reducionista impede a generalização para o discurso contínuo, onde os sinais se fundem, se modificam e se sobrepõem de forma fluida, conforme as regras fonológicas e prosódicas da língua.

Mais problemático ainda é o desprezo sistemático pelos parâmetros fonológicos nas arquiteturas atuais. Embora técnicas como OpenPose, MediaPipe ou HRNet forneçam estimativas detalhadas da posição das mãos e do rosto, raramente esses dados são mapeados para unidades linguísticas reconhecíveis — como configurações de mão ou movimentos contrastivos. Ao invés disso, os modelos operam com coordenadas cartesianas brutas ou vetores de pose normalizados, que capturam forma, mas não função linguística (ALMEIDA; LACERDA, 2024). Conseqüentemente, alterações fonológicas sutis — como a variação na orientação da palma que distingue “livro” de “carta” em Libras — muitas vezes são perdidas ou mal interpretadas, pois o sistema não foi projetado para reconhecer traços distintivos, mas sim padrões visuais correlacionais.

Essa limitação é agravada pela composição tendenciosa dos corpora utilizados para treinamento e validação. Estudos apontam que a maioria dos conjuntos de dados disponíveis — como o RWTH-PHOENIX-Weather (DGS), ASLLVD ou mesmo iniciativas brasileiras — foi coletada em contextos artificiais, com sinalizantes posicionados frontalmente diante de câmeras estáticas, sem interlocutores reais, e frequentemente com ouvintes fluentes em Libras, cuja produção pode diferir em aspectos prosódicos e fonológicos da de surdos nativos (DOGAN et al., 2022; SILVA et al., 2023). Além disso, há uma sub-representação de variações regionais, estilísticas e sociolinguísticas da Libras, o que introduz viés algorítmico e limita a aplicabilidade dos sistemas em contextos reais, diversos e dinâmicos.

Recentemente, algumas iniciativas têm buscado superar essas lacunas. Yin et al. (2023) propuseram um modelo multimodal que integra dados de pose, expressão facial e segmentação de mão com atenção diferenciada para parâmetros fonológicos, embora ainda de forma implícita. Já Almeida e Lacerda (2024) defenderam a necessidade de

anotação fonológica explícita em corpora de Libras, sugerindo um protocolo de codificação baseado nos parâmetros de Stokoe adaptado para aprendizado de máquina. Contudo, tais esforços permanecem isolados e raramente são adotados em larga escala pela comunidade de IA, que frequentemente prioriza métricas de acurácia técnica sobre validade linguística.

Assim, o estado da arte revela um campo em expansão tecnológica, mas estagnado em termos linguísticos. A ausência de diálogo com a fonologia visual não apenas compromete a eficácia dos sistemas, mas também reforça uma visão instrumental da língua de sinais, deslocando o foco do significado linguístico para o padrão visual. A seção seguinte discutirá como essa lacuna pode ser superada por meio de interfaces teórico-algorítmicas que reconheçam a estrutura fonológica como núcleo, e não periferia, do reconhecimento automático.

### **3. Interfaces entre fonologia visual e algoritmos: desafios e oportunidades**

A interseção entre fonologia visual e modelagem algorítmica representa um dos desafios mais instigantes — e urgentes — na fronteira entre linguística aplicada e inteligência artificial (IA). Enquanto a fonologia das línguas de sinais opera com base em parâmetros articulatorios simultâneos e espacializados, os modelos computacionais contemporâneos ainda carecem de arquiteturas capazes de capturar essa complexidade de forma linguisticamente significativa. Ainda assim, essa lacuna não é insuperável; ao contrário, abre espaço para inovações teórico-técnicas que podem redefinir tanto os fundamentos da linguística computacional quanto os princípios éticos da tecnologia inclusiva.

Um primeiro desafio reside na codificação dos parâmetros fonológicos em representações computacionais. Embora ferramentas de pose estimation (como MediaPipe ou OpenPose) forneçam coordenadas 2D ou 3D de juntas corporais, esses dados brutos não correspondem, por si só, a unidades linguísticas. A transição de “posição da mão” para “configuração de mão contrastiva” exige uma camada intermediária de análise fonológica estruturada. Nesse sentido, propostas recentes têm explorado esquemas de codificação baseados nos parâmetros de Stokoe (1960), adaptados para o contexto da Libras. Por exemplo, Almeida e Lacerda (2024) sugerem um sistema de feature vectors que mapeia configurações manuais segundo o sistema descritivo de Quadros e Karnopp (2004), enquanto expressões faciais são codificadas

com base em Action Units do modelo FACS (Facial Action Coding System), vinculadas a funções gramaticais específicas (ex.: sobrancelhas franzidas como marca de interrogação). Essa abordagem permite que redes neurais operem não com pixels ou coordenadas, mas com traços fonológicos discretos e contrastivos, aproximando a modelagem computacional da realidade linguística.

Essa codificação, contudo, só é viável na presença de corpora linguísticos multimodais rigorosamente anotados. Atualmente, a escassez de conjuntos de dados em Libras com anotação fonológica, prosódica e discursiva constitui um dos maiores entraves ao desenvolvimento de IA linguisticamente fundamentada. A maioria dos corpora existentes foca em transcrição lexical ou glosa, negligenciando os níveis subléxicos essenciais à fonologia visual (SILVA et al., 2023). Diante disso, faz-se urgente a construção de corpora multimodais que integrem, simultaneamente: (1) vídeo em alta resolução; (2) anotação de pose corporal e facial; (3) transcrição fonológica segmentada por parâmetros; e (4) glosa e/ou tradução para língua oral. Projetos como o Corpus Libras-UFSC (FERREIRA et al., 2022) e o Corpus Sinalizado Brasileiro (CSB) têm dado os primeiros passos nessa direção, mas ainda carecem de escala, diversidade sociolinguística e interoperabilidade com pipelines de IA. A colaboração entre linguistas, cientistas da computação e membros da comunidade surda é, portanto, não apenas desejável, mas condição necessária para a produção de dados que respeitem a língua em sua integralidade.

Além das questões técnicas, há uma dimensão semiótica fundamental que não pode ser negligenciada — especialmente em uma revista como a *Acta Semiótica e Linguística*, cujo escopo teórico articula linguagem, significação e prática social. A língua de sinais não é apenas um código fonológico: é um sistema semiótico que mobiliza iconicidade, indexicalidade e simbolização em interação dinâmica (SANDLER; LILLO-MARTIN, 2006). O movimento de um sinal pode não apenas indicar uma ação, mas também construir metáforas espaciais; a localização pode ser arbitrária ou, ao contrário, carregar referência discursiva por meio de pontos de ancoragem no espaço. Ignorar essa camada semiótica reduz o sinal a um rótulo classificatório, esvaziando-o de seu potencial expressivo e discursivo. Modelos de IA que pretendem “entender” línguas de sinais precisam, portanto, ir além da fonologia e incorporar princípios da semiótica visual — tal como proposto por autores como Kress

e van Leeuwen (2006) e, no contexto brasileiro, por Perlin (2018), que analisa as línguas de sinais como sistemas multimodais de construção de sentido.

Diante desses desafios, propõem-se as seguintes diretrizes para o desenvolvimento de IA linguisticamente informada:

1. Integração teórica: os pipelines de IA devem ser co-projetados com linguistas especializados em línguas de sinais, garantindo que os parâmetros fonológicos e as funções semióticas sejam explicitamente modelados.
2. Anotação fonológica padronizada: adotar esquemas de codificação fonológica interoperáveis (ex.: HamNoSys, SignWriting ou sistemas adaptados localmente) para permitir comparação e reutilização de dados.
3. Participação surda: envolver surdos nativos ou fluentes em todas as etapas — coleta, anotação, validação e avaliação —, respeitando o princípio da nada sobre nós sem nós.
4. Transparência algorítmica: documentar claramente quais aspectos linguísticos são (ou não) considerados pelos modelos, evitando a ilusão de neutralidade técnica.
5. Interdisciplinaridade: fomentar parcerias entre programas de pós-graduação em linguística, ciência da computação, design e estudos surdos, especialmente em instituições com infraestrutura de pesquisa, como a UFPA e outras universidades federais brasileiras.

Essas diretrizes não apenas melhoram a precisão técnica dos sistemas, mas também reafirmam o compromisso ético com a soberania linguística da comunidade surda. Ao tratar a fonologia visual não como um obstáculo técnico, mas como um recurso cognitivo e cultural, a IA pode tornar-se uma aliada na promoção da acessibilidade linguística com dignidade, rigor e respeito. A seção seguinte discutirá as implicações sociais, políticas e éticas dessa reorientação paradigmática.

#### **4. Implicações éticas, sociais e linguísticas**

O desenvolvimento de tecnologias baseadas em inteligência artificial (IA) para línguas de sinais não pode ser compreendido apenas como um desafio técnico ou computacional; trata-se, antes de tudo, de uma questão ética, social e linguística que repercute diretamente nos direitos humanos, na cidadania e na justiça linguística das pessoas surdas. Quando sistemas automatizados ignoram a estrutura fonológica e gramatical das línguas de sinais — tratando sinais como gestos isolados ou meros padrões visuais —, eles não apenas produzem traduções imprecisas, mas também

violam o direito à língua consagrado em instrumentos legais nacionais e internacionais. No Brasil, a Lei nº 10.436, de 2002, reconhece a Língua Brasileira de Sinais (Libras) como "meio legal de comunicação e expressão", atribuindo-lhe status de língua natural e garantindo seu uso em todos os espaços públicos e privados (BRASIL, 2002). A tradução automática deficiente ou distorcida, comum em sistemas que não consideram parâmetros fonológicos como orientação da palma ou expressões não manuais, pode levar a mal-entendidos graves em contextos sensíveis, como atendimentos médicos, processos judiciais ou emergências, comprometendo o acesso à informação, à saúde e à justiça — direitos fundamentais previstos na Convenção sobre os Direitos das Pessoas com Deficiência (ONU, 2006), ratificada pelo Brasil em 2008.

Esses riscos são amplificados pela exclusão sistemática da comunidade surda nos processos de concepção, desenvolvimento e validação de tecnologias de IA. Historicamente, soluções "para os surdos" têm sido projetadas por equipes majoritariamente ouvintes, sem a participação efetiva de surdos como coautores do conhecimento e da inovação (MIRANDA-GÁLVEZ; ALMEIDA, 2024). Essa lógica paternalista reproduz uma epistemologia da ausência, na qual o sujeito surdo é reduzido a um "usuário final" passivo, em vez de ser reconhecido como agente linguístico, cultural e tecnológico. Autores como Lacerda (2020) e Quadros (2021) destacam que a inclusão genuína exige co-projetar com, e não projetar para, a comunidade surda — o que implica remunerar expertise linguística, respeitar saberes surdos e garantir autonomia decisória nos processos de inovação. A participação surda não é um mero critério de "validação", mas uma condição ética e epistemológica para a produção de tecnologias que efetivamente respeitem a Libras em sua complexidade.

Além disso, a relação entre inovação tecnológica e políticas linguísticas deve ser reorientada. A Lei nº 10.436/2002 não apenas reconheceu a Libras, mas também estabeleceu diretrizes para sua promoção, ensino e disseminação. No entanto, a implementação dessas políticas frequentemente esbarra na falta de recursos humanos e tecnológicos qualificados. A IA, se bem direcionada, poderia ser uma aliada poderosa nesse processo — por exemplo, em plataformas de ensino adaptativo, sistemas de legendagem automática em Libras ou ferramentas de apoio à tradução em tempo real. Contudo, isso só será possível se os investimentos em tecnologia estiverem alinhados a uma política linguística soberana, que priorize a qualidade linguística, a diversidade regional da Libras e a formação de profissionais surdos em áreas STEM (Ciência,

Tecnologia, Engenharia e Matemática). Como observa Felipe (2012), a Libras não é um "recurso de acessibilidade", mas uma prática cultural constitutiva da identidade surda — e qualquer tecnologia que a instrumentalize sem esse reconhecimento corre o risco de reproduzir formas sutis de colonialidade linguística.

Nesse contexto, é possível estabelecer uma conexão relevante com a linha de pesquisa “Conservação e Construção Social do Meio Ambiente”, vinculada ao Programa de Pós-Graduação em Ciências e Meio Ambiente (PPGCMA/UFGA), ao qual a autora está associada. Embora tradicionalmente associada a ecossistemas naturais, a noção de “meio ambiente” pode — e deve — ser ampliada para incluir ambientes linguísticos e culturais como bens comuns a serem preservados e promovidos (HALL, 2017). A acessibilidade linguística, nessa perspectiva, não é um luxo técnico, mas um componente essencial de um ambiente social inclusivo e sustentável. Assim como a perda de biodiversidade ameaça o equilíbrio ecológico, a erosão linguística — provocada por tecnologias que normalizam formas distorcidas ou incompletas de línguas minoritárias — ameaça a diversidade cognitiva e cultural da humanidade. Nesse sentido, o desenvolvimento de IA linguisticamente informada para línguas de sinais pode ser compreendido como uma prática de inovação social sustentável, alinhada aos princípios de justiça ambiental ampliada, que inclui o direito a viver e comunicar-se em sua língua plenamente reconhecida.

Portanto, as implicações deste debate vão muito além da eficiência algorítmica: tocam o cerne da democracia linguística, da justiça cognitiva e da coexistência plural em sociedades hiperconectadas. Ignorar a fonologia visual não é apenas um erro técnico — é uma forma de silenciamento estrutural. Ao contrário, investir em tecnologias que respeitem e potencializem a estrutura linguística das línguas de sinais é contribuir para a construção de um ambiente social mais justo, inclusivo e verdadeiramente plurilíngue.

### **Considerações Finais/ Conclusão**

Este artigo partiu de uma pergunta central: como a fonologia das línguas de sinais é (ou não) considerada nos modelos algorítmicos atuais baseados em inteligência artificial? A análise desenvolvida ao longo das seções demonstra, de forma inequívoca, que a maioria dos sistemas contemporâneos opera com uma visão instrumentalizada e descontextualizada do sinal, relegando a estrutura fonológica — e, por extensão, a própria gramaticalidade da língua — a um plano secundário. Embora as tecnologias de

visão computacional tenham avançado significativamente na captura de movimentos corporais, elas frequentemente falham em reconhecer que, nas línguas de sinais, a forma é função — e que os parâmetros fonológicos (configuração da mão, localização, movimento, orientação e expressões não manuais) não são meros detalhes visuais, mas unidades distintas de significado, tal como os fonemas nas línguas orais (STOKOE, 1960; BRENTARI, 1998; QUADROS; KARNOPP, 2004).

A contribuição teórica deste trabalho reside em articular, de forma crítica e propositiva, dois campos frequentemente dissociados: a linguística de línguas de sinais e a ciência da computação aplicada à acessibilidade. Ao demonstrar que a fonologia visual não pode ser ignorada sem comprometer a própria validade linguística dos sistemas de IA, este artigo reforça a necessidade de uma fonologia computacional visual — um campo emergente que integre os princípios da linguística descritiva com as exigências da modelagem algorítmica (ALMEIDA; LACERDA, 2024; MIRANDA-GÁLVEZ; ALMEIDA, 2024). Ademais, ao vincular essa discussão à semiótica — campo central no escopo da *Acta Semiótica et Linguistica* —, evidencia-se que o sinal não é apenas uma forma articulatória, mas um ato de significação situado, que mobiliza iconicidade, espaço e corpo como recursos semióticos (SANDLER; LILLO-MARTIN, 2006; PERLIN, 2018).

Do ponto de vista social, este trabalho reafirma que acessibilidade linguística de qualidade é um direito, não um favor técnico. Sistemas que traduzem Libras de forma imprecisa ou estereotipada não apenas geram ruído comunicativo, mas também reforçam desigualdades estruturais e violam os direitos linguísticos consagrados na Lei nº 10.436/2002 e na Convenção sobre os Direitos das Pessoas com Deficiência (ONU, 2006). A inovação tecnológica só será verdadeiramente inclusiva se for co-construída com a comunidade surda, respeitando sua agência linguística e sua expertise cultural (LACERDA, 2020; QUADROS, 2021). Nesse sentido, a proposta aqui defendida — de tecnologias linguisticamente informadas e eticamente responsáveis — alinha-se também à perspectiva ampliada de “meio ambiente” adotada por linhas de pesquisa como a “Conservação e Construção Social do Meio Ambiente” (PPGCMA/UFPA), na qual a diversidade linguística é reconhecida como patrimônio sociocultural a ser protegido e promovido, assim como a biodiversidade natural (HALL, 2017).

Para pesquisas futuras, recomenda-se, em primeiro lugar, o desenvolvimento de corpora multimodais de Libras com anotação fonológica explícita, seguindo esquemas

padronizados e interoperáveis (ex.: HamNoSys ou SignWriting adaptado). Em segundo lugar, é urgente integrar teorias linguísticas diretamente nos pipelines de IA, de modo que parâmetros fonológicos sejam representados como features estruturais, não como ruído visual. Por fim, devem ser incentivadas parcerias interdisciplinares e interinstitucionais — envolvendo programas de pós-graduação em linguística, ciência da computação, estudos surdos e políticas públicas — que incluam, de forma remunerada e decisória, pesquisadores e ativistas surdos. Somente assim a inteligência artificial poderá transitar, de fato, entre gestos e algoritmos, sem perder de vista o que há de mais humano na linguagem: seu poder de significar, resistir e pertencer.

## Referências

- ALMEIDA, F. R.; LACERDA, C. B. F. Fonologia visual e aprendizado de máquina: desafios para a construção de corpora anotados em Libras. *Revista Brasileira de Linguística Aplicada*. Vol. 24, N. 1, p. 112–135, jan./mar. 2024. <https://doi.org/10.1590/1984-639820240301>
- BRASIL. Lei nº 10.436, de 24 de abril de 2002. Dispõe sobre a Língua Brasileira de Sinais – Libras e dá outras providências. *Diário Oficial da União*, Brasília, DF, 25 abr. 2002. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/2002/L10436.htm](https://www.planalto.gov.br/ccivil_03/leis/2002/L10436.htm). Acesso em: 28 dez. 2025.
- BRENTARI, D. *A Prosodic Model of Sign Language Phonology*. Cambridge: MIT Press, 1998.
- DOGAN, Ö. İ.; AYDIN, E.; KAYA, M. et al. A Survey on Sign Language Recognition: Datasets, Methods, and Challenges. *ACM Computing Surveys*. Vol. 55, N. 4, p. 1–37, mar. 2022. <https://doi.org/10.1145/3507905>
- FELIPE, T. A. O processo de reconhecimento da Libras como língua: implicações para a educação de surdos. In: QUADROS, R. M. (Org.). *Estudos Surdos II*. Petrópolis: Editora Arara Azul, 2012. p. 25–46.
- FERREIRA, A. C.; SILVA, M. L.; OLIVEIRA, P. R. Corpora sinalizados e inteligência artificial: avanços e desafios no contexto brasileiro. *SignWriting Journal*. Vol. 5, N. 1, p. 44–67, 2022. <https://doi.org/10.5281/zenodo.6543210>
- HALL, S. *A identidade cultural na pós-modernidade*. Tradução Tomaz Tadeu da Silva; Guaraci Corrêa. 13ª. Ed. Rio de Janeiro: Lamparina, 2017.
- HUENERFORTH, M.; MATTNER, M.; SCHNEIDER, D. et al. Towards Continuous Sign Language Recognition: Challenges in Linguistic Grounding. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

- Punta Cana: Association for Computational Linguistics, 2021. p. 456–471.  
<https://doi.org/10.18653/v1/2021.emnlp-main.35>
- KRESS, G.; VAN LEEUWEN, T. *Reading Images: The Grammar of Visual Design*. 2nd ed. London: Routledge, 2006.
- LACERDA, C. B. F. Inclusão digital e participação surda: desafios para a pesquisa e a inovação. *Revista Educação e Pesquisa*. Vol. 46, e211690, 2020.  
<https://doi.org/10.1590/s1678-4634202046211690>
- MIRANDA-GÁLVEZ, R.; ALMEIDA, F. R. Linguística de línguas de sinais e inteligência artificial: por uma ética do reconhecimento. *Revista Linguagem & Ensino*. Vol. 27, N. 2, p. 201–225, maio/ago. 2024. <https://doi.org/10.15210/rle.v27i2.33421>
- ORGANIZAÇÃO DAS NAÇÕES UNIDAS (ONU). *Convenção sobre os Direitos das Pessoas com Deficiência*. Nova York, 2006. Disponível em:  
<https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>. Acesso em: 28 dez. 2025.
- PERLIN, G. T. Semiologia do corpo surdo: linguagem, corpo e espaço. *Revista Línguas & Letras*. Vol. 19, N. 38, p. 12–29, jul./dez. 2018. <https://doi.org/10.5433/1982-1283.2018v19n38p12>
- PERLIN, G. T.; QUADROS, R. M. Fonologia da Libras: parâmetros e variação. In: QUADROS, R. M. (Org.). *Estudos Surdos I*. Petrópolis: Editora Arara Azul, 2008. p. 11–41.
- QUADROS, R. M. Pesquisa com comunidades surdas: ética, participação e coautoria. *Revista Brasileira de Educação Especial*. Vol. 27, p. 517–534, 2021.  
<https://doi.org/10.1590/1980-547020212702supl02>
- QUADROS, R. M.; KARNOPP, L. B. *Língua de Sinais Brasileira: estudos linguísticos*. Porto Alegre: Artmed, 2004.
- SANDLER, W.; LILLO-MARTIN, D. *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press, 2006.
- SILVA, A. V. R.; GUSTAVO, V. R.; ALMEIDA, F. R. Variação fonológica na Libras: implicações para o reconhecimento automático. *Revista Letras*. Vol. 103, p. 289–312, jan./abr. 2023. <https://doi.org/10.5380/letras.v103i0.89231>
- SILVA, J. C.; GÓES, M. C. R.; SOUZA, R. M. Expressões não manuais e prosódia na Libras: evidências experimentais. *Cadernos Cedes*. Vol. 42, N. 129, p. 301–320, set./dez. 2022. <https://doi.org/10.1590/0101-326220224212907>
- STOKOE, W. C. *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf*. Studies in Linguistics: Occasional Papers, n. 8. Buffalo: University of Buffalo, 1960.

YIN, H.; WANG, Y.; LIU, Z. et al. Multimodal Sign Language Recognition with Phonological-Aware Attention. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver: IEEE, 2023. p. 13456–13465. <https://doi.org/10.1109/CVPR52729.2023.01321>.