

DOCUMENTAR PARA INCLUIR, E NÃO APAGAR: CONSIDERAÇÕES METODOLÓGICAS PARA A CONSTRUÇÃO DE *CORPORA* DE SINAIS-TERMOS ESPECIALIZADOS EM LIBRAS

DOCUMENTING FOR INCLUSION, NOT ERASURE: METHODOLOGICAL CONSIDERATIONS FOR DEVELOPING SPECIALIZED SIGN-TERM *CORPORA* IN BRAZILIAN SIGN LANGUAGE

DOI:10.47677/glüks.v26i02.609

Recebido: 18/06/2026

Aprovado: 27/08/2026

MENDES, Rodrigo Geraldo¹

RESUMO: Este artigo tem o objetivo de apresentar o desenho metodológico desenvolvido em uma pesquisa de doutorado voltada à constituição de um *corpus* de sinais-termos matemáticos em Língua Brasileira de Sinais (Libras). Para isso, explicita uma arquitetura metodológica replicável para pesquisas que busquem documentar *corpora* terminológicos nessa língua, articulando pesquisa documental multimodal, Linguística de *Corpus*, anotação videográfica, descrição fonético-fonológica e avaliação comunitária por professores Surdos. Defende-se que a construção de um *corpus* em Libras não deve ser compreendida como criação aleatória de sinais nem como instrumento de padronização impositiva, mas como documentação situada de repertórios já em circulação. Ao focalizar as três primeiras fases da pesquisa, o artigo mostra como cada decisão metodológica contribui para preservar a materialidade visual-espacial da Libras, reconhecer a variação linguística, valorizar a identidade Surda e preparar bases para futuras práticas pedagógicas inclusivas.

PALAVRAS-CHAVE: *Corpus* em Libras; Sinais-termo; Metodologia; Identidade Surda; Inclusão.

1 Introdução

A constituição de *corpora* em línguas de sinais coloca desafios metodológicos específicos. Diferente dos fundamentados apenas em textos escritos, *corpora* sinalizados dependem do registro, da organização e da análise de dados visuais (Quadros 2016), frequentemente registrados em vídeos, imagens, sinalários, glossários multimodais, materiais

¹ Doutorando em Linguística Aplicada pela Universidade Estadual de Campinas (Unicamp). Professor Adjunto de Libras e Diversidade na Universidade Federal de Juiz de Fora (UFJF). E-mail: rodrigo.mendes@ufjf.br.

didáticos, provas traduzidas, redes sociais e práticas pedagógicas. No caso dos sinais-termo¹ matemáticos, esses desafios se intensificam porque o objeto não envolve apenas equivalências lexicais entre português e Língua Brasileira de Sinais (Libras), mas modos visualmente situados de produzir conceitos matemáticos em uma língua de modalidade visuoespacial.

Este artigo parte, de forma mais específica, de uma preocupação metodológica e epistemológica para documentar sinais-termo matemáticos em Libras sem tratar a variação como falha e sem reduzir a produção linguística Surda² a uma lista de sinais “corretos”. A proposta deste trabalho direciona-se por uma proposta de construção de *corpus* baseado na língua em circulação. Esse *corpus* não inventa sinais de forma aleatória, tampouco impõe uma nomenclatura única; ele localiza ocorrências, registra fontes, descreve variantes, organiza dados comparáveis e submete o material à avaliação de professores Surdos.

Nessa direção, a documentação linguística não deve ser confundida com práticas pontuais de criação artificial de sinais. Reunir professores Surdos em uma sala para propor sinais inexistentes e, em seguida, registrá-los, como se fossem representativos da Libras em uso, como já observado em projetos nacionais, não apreende, por si só, a complexidade linguística, identitária e cultural dessa língua. Sinais-termo emergem de práticas sociais, escolares, comunitárias e acadêmicas nas quais circulam, variam, estabilizam-se parcialmente e adquirem legitimidade pelo uso situado.

O artigo desdobra-se de uma pesquisa de doutorado organizada em quatro fases interdependentes: (1) levantamento documental dos sinais-termo matemáticos em Libras; (2) constituição do *corpus* e descrição fonético-fonológica das variantes; (3) avaliação dos sinais por professores Surdos da área de matemática; e (4) prática coletiva com professores Surdos da referida área sobre usos, sentidos e avaliações de sinais do *corpus*. Neste recorte, focalizam-se, em destaque, as três primeiras fases. A quarta fase não é descrita em detalhe neste recorte, mas permanece no horizonte da discussão, visto que as três primeiras constituem a base documental e descritiva que permitirá observar, posteriormente, como professores Surdos reconhecem, mobilizam, modificam ou negociam sinais matemáticos em situações de ensino-aprendizagem.

A arquitetura metodológica apresentada pode ser aplicada a distintos domínios terminológicos em Libras, incluindo ciências naturais, saúde, tecnologias e humanidades. O artigo, portanto, concentra-se em discutir as escolhas metodológicas que sustentam a

elaboração de um *corpus* de sinais-termo em Libras, destacando de que forma tal processo pode promover o aprimoramento da documentação, a ampliação da inclusão e o fortalecimento da identidade Surda.

É importante destacar que essa opção se torna particularmente relevante para pesquisas em Libras porque os dados provenientes de línguas sinalizadas não se deixam capturar integralmente por transcrições verbais. Nesse sentido, a forma do sinal, sua temporalidade, a simultaneidade entre mãos, expressões faciais, sua realização no espaço e a origem social de cada ocorrência, precisam ser preservadas como dimensões de análise. Por isso, a replicabilidade aqui defendida não corresponde a repetir mecanicamente todo um protocolo, mas a sustentar uma lógica de documentação que possa ser recontextualizada para outros propósitos e objetivos de pesquisa.

Para que essa arquitetura seja legível como proposta replicável, o artigo privilegia a explicitação dos encadeamentos entre decisões, instrumentos e produtos metodológicos. Assim, cada fase é apresentada a partir de quatro perguntas: (a) quais tipos de dados são produzidos? (b) por que esses dados são necessários? (c) como são organizados? d) quais papéis cumprem na etapa seguinte? Essa forma de exposição evita que a metodologia apareça como inventário de ferramentas e permite compreendê-la como percurso de construção de evidências.

2 Linguagem, identidade surda e inclusão: o *corpus* como mediação

2.1 Fundamentos teóricos: linguística Aplicada, multimodalidade e epistemologias Surdas

A proposta metodológica apresentada neste artigo se inscreve em uma concepção de Linguística Aplicada que não se limita à aplicação de teorias linguísticas a problemas educacionais previamente definidos. Conforme Moita Lopes (2006, 2013), a Linguística Aplicada contemporânea se constitui como campo INdisciplinar, preocupado com problemas socialmente situados, atravessados por linguagem, poder, identidade e desigualdade. Ao mobilizar essa perspectiva, a proposta aqui discutida não é concebida como um instrumento neutro de coleta de sinais, mas como uma resposta metodológica a um problema de

linguagem situado, marcado pela documentação insuficiente, pela descrição limitada e pela circulação ainda restrita de repertórios matemáticos sistematizados em Libras.

Ao invés de tomar os sinais-termo como meras equivalências conceituais estáticas, a pesquisa os compreende como repertórios semióticos em circulação. Apoiar-se na Linguística Aplicada justifica-se porque a ausência ou a padronização compulsória de terminologias em Libras constituem um problema social de linguagem atravessado por relações de poder, exclusão acadêmica e lutas identitárias Surdas. Um *Corpus* de sinais matemáticos, assim, não é um mero inventário técnico, mas uma intervenção política e epistemológica sobre as práticas de linguagem que vão mediar o ensino de ciências.

Essa perspectiva também justifica a centralidade atribuída à abordagem êmica. Em Pike (1967), a distinção entre perspectivas êmica e ética aponta para diferentes modos de produzir conhecimento. Enquanto a perspectiva ética descreve fenômenos a partir de categorias externas ao grupo investigado, a perspectiva êmica busca compreender os sentidos com base nas categorias, experiências e práticas dos próprios participantes. Neste artigo, a abordagem êmica se materializa tanto na trajetória do pesquisador Surdo quanto na participação de professores Surdos na avaliação dos sinais-termo, deslocando a comunidade surda da posição de objeto de pesquisa para a posição de agentes de produção e validação do conhecimento.

A fundamentação teórica do artigo dialoga, ainda, com os Estudos Surdos e com a compreensão sociocultural da surdez. Autores como Padden e Humphries (2000), Ladd (2003), Strobel (2008) e Hauser *et al.* (2010) contribuem para reposicioná-la de uma matriz clínico-deficitária para uma matriz linguística, cultural e epistemológica. Nessa perspectiva, a Libras não é apenas um recurso de acessibilidade ou meio auxiliar de comunicação. Ela é língua de produção de mundo, de pertencimento, de memória comunitária e de construção de conhecimento. Por isso, documentar sinais-termo matemáticos em nessa língua exige reconhecer que tais sinais participam de repertórios culturais, pedagógicos e identitários, e não apenas de listas lexicais ou equivalências terminológicas.

Essa discussão se articula diretamente à noção de epistemologias Surdas. Hauser *et al.* (2010) argumentam que sujeitos Surdos produzem modos próprios de conhecer, fortemente atravessados pela experiência visual, pela língua de sinais e pelas práticas comunitárias de circulação do saber. No campo educacional, essa perspectiva impede que a inclusão seja

reduzida à presença física do estudante Surdo em espaços majoritariamente ouvintes. A inclusão precisa envolver condições linguísticas e epistemológicas para que o conhecimento seja produzido, negociado e avaliado em língua de sinais. Desse ponto de vista, a constituição de um *corpus* de sinais-termo em Libras contribui para tornar visíveis formas Surdas de nomear, organizar e ensinar conceitos mais especializados.

A materialidade visuoespacial da Libras também aproxima o artigo dos estudos sobre multimodalidade e multiletramentos. Kress e van Leeuwen (2001), Kress (2010) e Jewitt (2008) demonstram que os sentidos não são produzidos exclusivamente por meio da escrita verbal, mas pela articulação entre diferentes modos semióticos, como imagem, gesto, corpo, espaço, movimento e recursos audiovisuais. Kalantzis, Cope e Pinheiro (2020), em diálogo com a pedagogia dos multiletramentos, enfatizam que práticas contemporâneas de linguagem exigem atenção à multiplicidade de modos, mídias, culturas e formas de participação. Esse referencial é decisivo para pesquisas em Libras, pois o sinal não pode ser adequadamente descrito se for reduzido a uma palavra em português ou a uma glosa isolada.

No campo específico da documentação de línguas de sinais, Leite e Quadros (2014), Quadros (2015, 2016, 2017), Stumpf *et al.* (2020), Schembri *et al.* (2013) e Crasborn e Fiévez (2021) evidenciam que *corpora* sinalizados dependem de procedimentos capazes de preservar a temporalidade, a simultaneidade e a espacialidade dos dados visuais. A organização de vídeos, trilhas de anotação, glosas, metadados e parâmetros fonético-fonológicos não constitui, portanto, uma exigência meramente operacional. Trata-se de condição metodológica indispensável para que a Libras seja analisada em sua forma própria, sem ser subordinada a modelos analíticos desenvolvidos para as línguas orais-auditivas.

2.2 Variação Linguística, Identidade Surda e Inclusão

A discussão sobre sinais-termo especializados em Libras recorrentemente é atravessada por uma tensão entre padronização e variação terminológica. Em campos como a matemática, marcados por currículos, exames nacionais e terminologias estabilizadas em português, pode parecer desejável produzir uma lista única de sinais para cada conceito. Essa expectativa, contudo, precisa ser problematizada. Amparando-se em Skliar (1998), Lodi, Harrison e Campos (2015) torna-se possível afirmar que quando a padronização é tomada

como apagamento da heterogeneidade, pode converter a diversidade linguística da Libras em problema a ser corrigido e, com isso, enfraquecer formas comunitárias de produção de conhecimento.

Essa tensão entre padronização e variação já vem sendo discutida. Quadros e Karnopp (2004), ao descreverem aspectos fonológicos, morfológicos e sintáticos dessa língua, contribuem para compreendê-la como língua natural, dotada de regularidades próprias, como também atravessada por usos sociais diversos. Nessa mesma direção, Castro Júnior (2011, 2014), ao tratar da variação linguística na Libras, especialmente no léxico, problematiza a ideia de uma língua homogênea e evidencia que diferentes comunidades, regiões e grupos de usuários produzem e legitimam formas distintas de sinalização. Xavier e Barbosa (2014), ao focalizarem a variação fonológica, também demonstram que diferenças na realização dos sinais não devem ser tomadas como desvios ou inadequações, mas como parte do funcionamento linguístico da Libras. Desse modo, a documentação de sinais-termo matemáticos precisa ser compreendida como processo de registro, descrição e análise de repertórios em circulação, e não como a busca por uma forma única, definitiva e universalmente válida para cada conceito.

A variação terminológica, nesse sentido, não constitui apenas um dado identitário e cultural, mas desempenha uma função pedagógica e acadêmica fundamental. Em contextos de alta padronização, como avaliações escolares, exames nacionais e processos seletivos, estudantes Surdos podem deparar-se com um termo matemático em português escrito, ter aprendido o conceito por meio de uma variante regional e ser acompanhados por intérpretes que utilizam outra forma sinalizada. Nesses casos, registrar as variações e apresentá-las aos usuários da língua não geraria confusão semântica nem produziria ambiguidades. Ao contrário, amplia suas possibilidades de reconhecimento conceitual e de trânsito entre repertórios locais, comunitários e institucionais.

Por isso, o *corpus* é compreendido como uma mediação entre três dimensões: (1) a linguagem, porque registra a Libras em sua materialidade visuoespacial; (2) a identidade surda, porque reconhece a legitimidade dos repertórios produzidos por sujeitos Surdos em práticas sociais situadas; e (3) a inclusão, porque oferece instrumentos para que professores, intérpretes e estudantes negociem sentidos matemáticos sem depender exclusivamente de

traduções improvisadas ou de listas normativas. Nessa perspectiva, documentar significa incluir sem apagar a diversidade linguística.

Não se trata apenas de perguntar qual é o sinal “correto” para determinado termo em português, mas de investigar quais sinais circulam, em quais materiais são registrados, por quais parâmetros fonético-fonológicos se caracterizam, quem os reconhece, quem os utiliza e em que condições eles podem apoiar práticas pedagógicas. Essa mudança de perspectiva permite articular documentação linguística, inclusão educacional e identidade surda sem subordinar uma dessas dimensão às demais.

Dessa forma, a manutenção da identidade surda é reconhecida como princípio transversal da pesquisa, visto que atua não como um tópico isolado, mas como uma diretriz central que atravessa, orienta e perpassa todas as fases do desenho metodológico. Longe de ser um tema externo à organização dos dados, ela funciona como um critério que rege a seleção das fontes, a preservação da variação e a participação de professores Surdos na avaliação dos sinais. Um *corpus* que apaga a origem comunitária dos repertórios ou que transforma variantes em desvios acaba por produzir uma inclusão de caráter assimilacionista. Em direção distinta, a proposta aqui delineada entende que verdadeira inclusão exige documentar as formas pelas quais a comunidade surda já nomeia, visualiza, ensina e negocia conceitos especializados.

3 Arquitetura geral da pesquisa e sua replicabilidade

A pesquisa foi organizada em um fluxo metodológico em camadas em que cada etapa alimentou a fase subsequente. Um levantamento documental identificou sinais em circulação, a constituição do *corpus* transformou ocorrências em dados comparáveis, e a avaliação por professores Surdos agregou uma leitura comunitária dos sinais-termo. Na investigação que deu origem a este estudo, uma quarta fase foi desenvolvida e pautou-se em observar a mobilização desses repertórios em práticas linguístico-pedagógicas pelos professores Surdos.

Fase	Finalidade	Produto metodológico	Potencial de replicação
Fase 1 Levantamento Documental	Localizar repertórios já registrados em diversas fontes.	Mapa de fontes e conjunto inicial de sinais-termo.	Aplicável a áreas que possua registros multimodais de sinais em circulação.
Fase 2 Constituição e Descrição do <i>Corpus</i>	Organizar ocorrências, indexar variantes, anotar vídeos e descrever parâmetros fonético-fonológicos.	Base documental, fichas, glosas, vídeos anotados e descrições comparáveis.	Replicável por meio de planilhas, uso de <i>Softwares</i> de notação de vídeo por trilhas, convenções de glosa e parâmetros fonético-fonológicos.
Fase 3 Avaliação Comunitária	Submeter termos e suas variantes à avaliação de professores Surdos.	Avaliação com base no reconhecimento, uso, adequação e circulação dos sinais.	Replicável com a participação de Surdos da área investigada.

Quadro² 1 - Arquitetura metodológica das fases focalizadas no artigo

Esse desenho recusa duas soluções simplificadoras. A primeira seria partir da hipótese de ausência, como se a comunidade surda não produzisse repertórios especializados até que uma pesquisa acadêmica os criasse. A segunda seria transformar o *corpus* em instrumento normativo, destinado a selecionar uma única forma legítima para cada termo. O percurso aqui proposto toma outra direção, visto que seu propósito está em mapear repertórios existentes, descrever variações e criar condições para que sejam avaliadas comparativa e pedagogicamente.

Para replicações em estudos de outras áreas, esse desenho sugere um princípio simples: antes de propor intervenção, elaboração de material didático ou discussão normativa, é necessário conhecer como os sinais circulam. Isso significa localizar fontes, registrar metadados, distinguir ocorrências de variantes, descrever a forma sinalizada e submeter o repertório à avaliação de sujeitos com experiência na área. O método, portanto, contribui tanto para produzir conhecimento linguístico quanto para qualificar decisões pedagógicas e terminológicas.

Do ponto de vista da arquitetura geral, as fases podem ser compreendidas como um encadeamento de produtos metodológicos. A fase documental objetiva construir um inventário rastreável de fontes e ocorrências; a fase de constituição do *corpus* produz uma base indexada, anotada e descrita fonético-fonologicamente; a fase de avaliação comunitária propõe uma atividade situada sobre reconhecimento, uso e adequação dos sinais; e a quarta fase, não desenvolvida neste artigo, teve como objetivo observar como esses repertórios são mobilizados em contexto bilíngues (Libras /Língua Portuguesa). Essa sequência sustenta a

² Todos os quadros e figuras foram elaboradas pelo autor.

progressão metodológica do estudo e impede que cada procedimento seja interpretado de modo isolado.

3.1 Fase 1: Pesquisa Documental Multimodal

A primeira fase envolveu o levantamento de materiais que registram sinais-termo matemáticos em Libras em circulação, incluindo fontes de reconhecimento institucional e acadêmico e glossários impressos, vídeos digitais e sinalários produzidos por pesquisadores, professores ou instituições em geral. O foco foi identificar todos os meios reais de ocorrência desses sinais e não apenas os de reconhecimento acadêmico. Considerando que os sinais-termo podem existir e circular na comunidade surda antes de serem documentados por pesquisas ou incorporados a glossários especializados, o levantamento contemplou múltiplos meios de ocorrência. Assim, foram consideradas tanto fontes de reconhecimento institucional e acadêmico quanto espaços de circulação social da Libras, incluindo materiais didáticos, plataformas digitais, produções audiovisuais, redes sociais, práticas educacionais e o uso cotidiano da comunidade surda. Essa opção metodológica buscou identificar os sinais efetivamente em uso, evitando restringir o corpus apenas às formas previamente legitimadas pela literatura científica.

Essa escolha se apoia em uma concepção ampliada de pesquisa documental. Yonaha (2024), ao discutir a pesquisa documental na Linguística Aplicada, observa que os materiais de análise podem incluir “leis, manuais, jornais, fotografias, gravações, entre outros” (p. 5), o que desloca o documento para além do texto verbal escrito e amplia o escopo do que pode ser tomado como objeto de investigação. Essa ampliação é decisiva em pesquisas sobre Libras, pois muitos fenômenos linguísticos da comunidade surda se materializam em registros visuais, videográficos e multimodais. Desse modo, vídeos em plataformas de armazenamento online, fotografias, glossários pictográficos e outros materiais digitais não funcionam como meros complementos ilustrativos, mas como documentos centrais para a análise da língua em uso.

A fase documental, portanto, cumpre duas funções: (a) empírica, com a identificação dos sinais-termo já disponíveis e suas fontes e (b) epistemológica, ao reconhecer que a produção linguística em Libras ocorre em múltiplos espaços e não apenas em materiais acadêmicos formais. Essa decisão impede que o *corpus* seja construído a partir de uma única

fonte de autoridade e preserva a heterogeneidade dos repertórios em circulação. Também é importante distinguir fonte documental e unidade de análise. Um glossário, uma videoaula ou uma postagem em rede social constitui uma fonte, uma vez que o sinal-termo localizado nesses materiais constitui uma ocorrência. A comparação entre ocorrências é, portanto, o que torna possível a identificação de variantes. Essa distinção evita que o *corpus* confunda material de origem com dado linguístico e ajuda a preservar a rastreabilidade entre o sinal analisado e seu contexto de circulação.

Na prática, essa fase exige critérios bem delimitados de busca e de inclusão. Em um estudo replicável, recomenda-se registrar os descritores utilizados, as plataformas consultadas, o período de acesso, o tipo de material encontrado, a autoria ou instituição responsável, a forma de representação do sinal e as condições de acesso ao dado. Esses elementos permitem que o levantamento seja auditável e que outros pesquisadores compreendam por que determinados materiais foram incluídos ou excluídos. A título de exemplificação, o quadro 2 ilustra como se deu essa organização na pesquisa da qual deriva a metodologia aqui descrita.

Fonte	Catálogo consultado	Meio de publicação	Modalidade de representação	Nº de sinais
Atayde (2019)	Glossário: 144 sinais	Impresso	Pictográfica	115
Lobato (2015)	Glossário: 90 sinais (27 por datilologia)	Impresso	Fotográfica	63
Martins (2019)	Glossário: 32 sinais	Impresso + <i>YouTube</i>	Fotográfica e videográfica	32
Carvalho (2017)	Glossário: 266 sinais + 50 novos sinais (atualizado no ano 2025)	Blog educacional + <i>YouTube</i>	Videográfica	316
D'Azevedo (2019)	Glossário matemática Libras	<i>Online</i>	Videográfica	30
INEP (2022) ³	Glossário Enem Libras: Matemática e suas Tecnologias	Online institucional	Videográfica	110
Zanubia Dada (2013) ^{4,5}	Sinalários de matemática em Libras	<i>YouTube</i>	Videográfica	106 / 56 / 34
Instituto Phala (2013; 2017; 2022) ⁶	Sinalários de matemática, medidas e formas geométricas	<i>YouTube</i>	Videográfica	19 / 16 / 16

Quadro 2 - Síntese dos catálogos contemplados, modalidades e números de sinais

Como resultado dessa fase, delimitou-se um conjunto inicial de sinais-termo matemáticos e de fontes documentais. Para fins de replicação, a etapa pode ser organizada em

³ Disponível em: https://www.youtube.com/inep_oficial%22%20/h. Acesso em: 15 set. 2024.

⁴ Disponível em: <https://www.youtube.com/@Dadaza100>. Acesso em: 20 jul. 2021.

⁵ Disponível em: <https://www.youtube.com/@zanubiadada7366>. Acesso em: 20 jul. 2021.

⁶ Disponível em: <https://www.youtube.com/user/institutophala/featured>. Acesso em: 20 jul. 2021.

quatro movimentos: definição dos descritores de busca; seleção de fontes acadêmicas e comunitárias; registro do tipo de material e de sua modalidade semiótica; e consolidação de um inventário inicial de ocorrências. O aspecto decisivo é manter a rastreabilidade para que cada sinal permaneça ligado à fonte em que foi encontrado.

Em *corpora* de outras áreas, como biologia, química, informática ou saúde, a mesma lógica pode ser mantida. O pesquisador pode alterar os descritores e as fontes, mas conservar o princípio de documentação multimodal: registrar materiais escritos, visuais e videográficos; diferenciar produções institucionais e comunitárias; e reconhecer que repertórios terminológicos especializados em Libras circulam em espaços formais e informais de aprendizagem.

Essa organização colabora com o leitor porque evidencia que a pesquisa documental não é uma etapa preparatória menor. Ela define a qualidade do *corpus* que será produzido nas fases seguintes. Se as fontes forem pouco variadas, o *corpus* tenderá a reproduzir uma imagem estreita da língua, se os metadados forem insuficientes, as variantes perderão sua história de circulação e se os vídeos não forem preservados, a descrição fonético-fonológica ficará dependente de registros secundários. Por isso, a fase 1 deve ser compreendida como o fundamento metodológico de todo o percurso.

Na metodologia em que se baseia este artigo, foram levantados 560 termos matemáticos e 913 ocorrências de sinais. Para análise inicial da referida pesquisa, foram priorizados termos com maior número de ocorrências e/ou maior diversidade de variantes, uma escolha metodológica feita diante do alto número de termos. Alguns conceitos matemáticos apresentam maior estabilização, enquanto outros, como quadrado, triângulo, cubo, cilindro, retângulo, fração e potência, demonstram que a circulação dos sinais não é homogênea. Eles revelam repertórios concorrentes, regionais ou dependentes de contextos específicos de ensino e documentação.

Os termos POTÊNCIA e CUBO foram selecionados neste artigo como modelos de indexação para exemplificar os movimentos da fase seguinte, uma vez que permitem visualizar dois movimentos importantes, os quais de um lado temos a variação em torno de um conceito matemático abstrato e, de outro, a variação em torno de um conceito geométrico de forte apelo visual.

3.2 Fase 2: Constituição da base documental e descrição fonético-fonológica

Na segunda fase, o foco desloca-se da localização das fontes e quantificação dos termos para a transformação das ocorrências em *corpus*. Essa etapa foi ancorada nos princípios da Linguística de *Corpus* e nos protocolos de documentação de línguas de sinais. Com Berber Sardinha (2004) e Biber, Conrad e Reppen (1998), compreende-se a Linguística de *Corpus* como abordagem empírica voltada à análise de dados autênticos. No caso da Libras, essa abordagem permite observar padrões, variantes e recorrências que dificilmente seriam percebidos por introspecção ou por listagens isoladas.

A adoção da Linguística de *Corpus*, contudo, não é apenas técnica. Ela assume valor epistemológico, pois desloca a análise de modelos idealizados para práticas efetivas de uso. Flores e Rebechi (2023) lembram que pesquisas que articulam Linguística de *Corpus* e línguas de sinais ainda são escassas, em parte pelas dificuldades de transcrição e anotação de dados visuais. Justamente por isso, a constituição de *corpora* sinalizados torna-se estratégica, pois oferece instrumentos para registrar a língua visual em uso, tornar a variação observável e produzir descrições comparáveis.

Em diálogo com esse campo, a documentação de *corpora* em línguas de sinais demanda atenção específica à relação entre dado primário e representação analítica. Schembri *et al.* (2013), ao discutirem a construção do *British Sign Language Corpus*, mostram que *corpora* sinalizados exigem decisões sistemáticas sobre gravação, segmentação, metadados, anotação e acesso aos vídeos. Crasborn e Fiévez (2021) também destacam que a pesquisa com *corpora* de línguas de sinais precisa lidar com a complexidade de dados multimodais, nos quais a anotação escrita não substitui o vídeo, mas cria camadas de leitura sobre ele.

A operacionalização da fase 2 se deu em três camadas interdependentes, conforme sintetiza o quadro a seguir:

Fase 2	Operacionalização	Função
Camada 1	Preenchimento de fichas documentadas armazenadas no <i>Google Sheets</i>	Rastrear a procedência do dado
Camada 2	Armazenamento de vídeos e anotação de glosas no <i>software ELAN</i>	Registrar como a ocorrência dos sinais-termo se realizam temporalmente no vídeo
Camada 3	Elaboração de fichas com descrição fonético-fonológica dos termos	Apresentar e organizar cada termo no intuito de possibilitar comparações entre suas variantes

Quadro 3 - Síntese das camadas que constituem a fase 2

A força do desenho metodológico está na articulação dessas camadas de forma que nenhuma delas, isoladamente, daria conta de sustentar um *corpus* robusto. Na camada 1, uma base documental foi estruturada em planilhas no *Google Sheets* em forma de fichas (vide figura 1). Essa escolha não deve ser entendida como simples armazenamento, visto que as planilhas funcionam como núcleo de rastreabilidade do *corpus*. Nas fichas há indicação de onde cada sinal foi localizado, em que modalidade foi registrado, a que termo se vincula, qual variante representa e quais arquivos ou vídeos dão acesso à ocorrência. Registrar se determinado sinal aparece em glossário institucional, vídeo de professor Surdo, material acadêmico ou sinalário comunitário significa reconhecer a trajetória social do repertório documentado.

ID-Termo	(identificador do termo)	
Termo	(palavra com significado especializado, usada em uma área de conhecimento específico)	
Ficha ID	(identificador da ficha fonológica do termo)	
Nº de ocorrências de Sinais	(número de ocorrências de sinais do termo)	
Nº de Sinais Diferentes	(número de sinais distintos do termo)	
Ocorrência nº	xxx	(enumerar cada ocorrência em ordem crescente)
ID Autor	(identificador do autor da fonte)	
Tipo de publicação	(formato de publicação da fonte)	
Ano de publicação	(ano de publicação da fonte)	
Forma de representação	(modalidade visual do sinal compilado pela fonte)	
Acesso da ficha anotada	(informação sobre a ficha anotada pelo autor, caso houver)	
Local do Sinal identificado	(informação de local de acesso ao sinal publicado na fonte)	
Acesso ao Sinal identificado	(localização de acesso ao sinal do termo)	
Registro <i>ELAN</i>		
ID-Sinal	(identificador do sinal)	
Glosa ID	(identificador da glosa do sinal anotado no <i>ELAN</i>)	

Figura 1: Ficha documentada de termo com sinais-termo da Libras

Essa memória é indispensável quando se trabalha com dados multimodais em ambientes digitais, pois *links* podem sair do ar, vídeos podem ser editados e materiais podem circular sem estabilidade editorial. Ao registrar metadados e cópias de acesso, o pesquisador reduz a fragilidade do *corpus* e cria condições para revisões posteriores. Esse cuidado é parte da ética documental da pesquisa em Libras, especialmente quando o repertório provém de práticas comunitárias e pedagógicas.

Em termos de replicabilidade, a planilha deve funcionar como matriz de controle dos metadados e memória do *corpus* (Cf. Berber Sardinha, 2004). Campos como Identificador do Termo (ID-Termo), termo em português, número de ocorrências, número de sinais diferentes, fonte, ano, tipo de publicação, modalidade de representação, *link* de acesso, Identificador do

Sinal (ID-Sinal) e glosa impedem que o sinal seja tratado como item isolado. Isso ocorre justamente pois vincula cada ocorrência a uma cadeia de informações, como termo, fonte, publicação, modalidade de representação, acesso ao vídeo ou imagem, variante e glosa. Para outros estudos, essa estrutura pode ser replicada com adaptações que atendam aos objetivos do pesquisador desde que se mantenha sua função principal de rastreabilidade.

Esse procedimento torna a metodologia transferível para outros *corpora*. Em uma pesquisa sobre sinais-termo de saúde, por exemplo, a ficha poderia manter a mesma lógica, alterando apenas o campo terminológico e os tipos de fonte. Em uma pesquisa sobre sinais de tecnologia, poderiam ser acrescentados campos sobre circulação em redes sociais ou comunidades profissionais. A estrutura é relativamente estável, pois sua aplicação permanece sensível ao propósito investigativo de outras pesquisas.

A ficha também permite distinguir documentação, anotação e interpretação. A documentação registra a existência e a fonte da ocorrência; a anotação segmenta e nomeia a unidade no vídeo; a interpretação descreve parâmetros e compara variantes. Quando essas operações ficam separadas, o leitor consegue acompanhar o percurso que transforma um sinal localizado em uma fonte digital ou impressa em dado linguístico analisável.

A camada 2, cujo foco é armazenar os vídeos e proceder à anotação de glosas no *software ELAN*¹, faz parte do procedimento metodológico de grande relevância para registrar distintas recorrências de variações fonético-fonológicas e lexicais dos termos. É importante salientar que um termo pode apresentar muitas ocorrências, mas poucas variantes, ao passo que outro pode ter menor frequência e maior diversidade formal. Essa distinção orienta análises posteriores sobre estabilização, dispersão e circulação dos sinais.

A escolha por trabalhar com os dados videográficos integrando-os ao *ELAN* se deve à sua compatibilidade com materiais audiovisuais e à capacidade de organizar anotações em trilhas temporais. No contexto da Libras, o *ELAN* possibilita manter a simultaneidade e a espacialidade dos sinais, criando camadas específicas para mão direita, mão esquerda, comentários do transcritor, tradução para o português e glosa grafadas em maiúsculo. Dessa forma, a ferramenta não só torna a transcrição possível, mas também impede que a visualidade dos dados seja limitada a uma simples equivalência verbal.

Esse ponto é central para a replicação do método. Em *corpora* sinalizados, a transcrição deve ser entendida como representação analítica parcial, não como substituição do dado visual. O

vídeo continua sendo a fonte primária, enquanto a anotação organiza pontos de entrada para leitura, comparação e descrição. Essa relação entre vídeo e anotação protege a pesquisa contra o risco de verbalizar excessivamente a Libras.

Na pesquisa em que essa metodologia foi desenhada e utilizada, a anotação no *ELAN*² foi organizada por trilhas mínimas, entre elas mão dominante, mão não dominante, e observações do anotador. Essa divisão permitiu observar a estrutura simultânea da Libras sem reduzir o sinal a uma glosa única. A glosa cumpre uma função de indexação e busca, mas não substitui a forma sinalizada. Por essa razão, permanece vinculada ao vídeo e às camadas descritivas posteriores.

Com base no *Signbank* da Libras (Stumpf *et al.*, 2020), no manual de transcrição da Libras (Quadros, 2015) e no Catálogo Nacional das Configurações de Mão do INES/MEC³, elaboramos um Manual de Categorização no âmbito do projeto em que foram definidos códigos para os principais parâmetros de descrição dos sinais. O manual sintetiza as categorias necessárias para que cada ocorrência seja descrita de forma estável, comparável e sensível à modalidade visual-espacial da Libras, acrescentando rigor à descrição fonético-fonológica dos termos do *corpus*.

Código/campo	Função no <i>corpus</i>
ID-Sinal	Número sequencial atribuído no corpus, independente de identificadores externos
Glosa ID	Identificador anotado no <i>ELAN</i> , em letras maiúsculas, conforme convenções apresentadas pelo tutorial de transcrição da Libras (Quadros, 2015)
LAT-Mao	Registro da quantidade de mãos e lateralidade: uma mão, duas mãos simétricas, assimétricas, não espelhadas ou ausentes
CM	Configuração de mão, descrita com base no catálogo do INES/MEC
OR	Orientação da mão, incluindo posição da palma/dorso e mudanças de orientação do pulso
LOC	Locação ou ponto de articulação, incluindo corpo, cabeça, extremidades e espaço neutro
MOV	Movimento, considerando forma, direção, contato, alternância e repetição
ENM	Expressões não manuais de face, olhos, cabeça, tronco, boca e funções sintático-discursivas
Observação	Campo livre para informações contextuais ou descritivas relevantes

Quadro 4 - Síntese dos códigos para descrição fonético-fonológica

(Fonte: elaborado pelo autor com base no Manual de Categorização para Anotação Fonético-Fonológica de *Corpus*)

O manual com códigos para descrição fonético-fonológica sintetizado acima, portanto, não é um apêndice técnico isolado. Ele opera como instrumento de padronização metodológica interna, sem padronizar a língua. Sua função é estabilizar a descrição, não

apagar a variação. Ao definir campos como ID-Sinal, Glosa ID, LAT-Mao, CM, OR, LOC, MOV e ENM, o manual torna possível comparar variantes de um mesmo termo e sinais associados a termos diferentes, preservando a singularidade de cada ocorrência.

A vantagem desse modelo está justamente na abertura analítica que ele proporciona. Campos controlados favorecem comparação entre sinais e reduzem variações de anotação entre pesquisadores, de forma ampliada; campos abertos, por sua vez, preservam particularidades de ocorrência, comentários de uso, dúvidas de transcrição e aspectos contextuais sendo especialmente produtivos para línguas de sinais, nas quais a descrição precisa ser precisa sem perder a riqueza multimodal do dado. A síntese dos códigos do manual torna a descrição mais econômica e mais comparável. Abaixo, as figuras 2 e 3 exemplificam como essa descrição fonético-fonológico do termo foi realizada no âmbito da pesquisa.

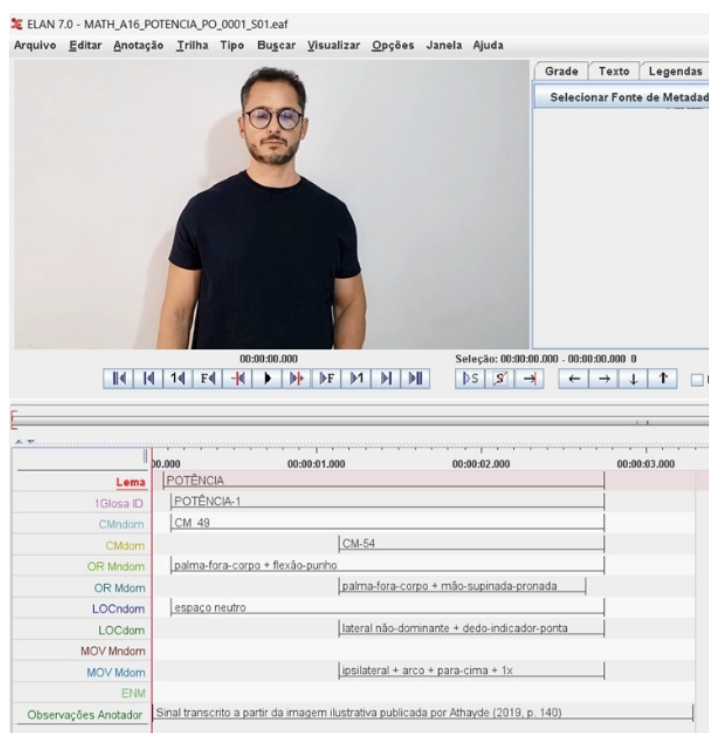


Figura 2: Captura de tela de anotação fonético-fonológica do sinal-termo POTÊNCIA do ELAN

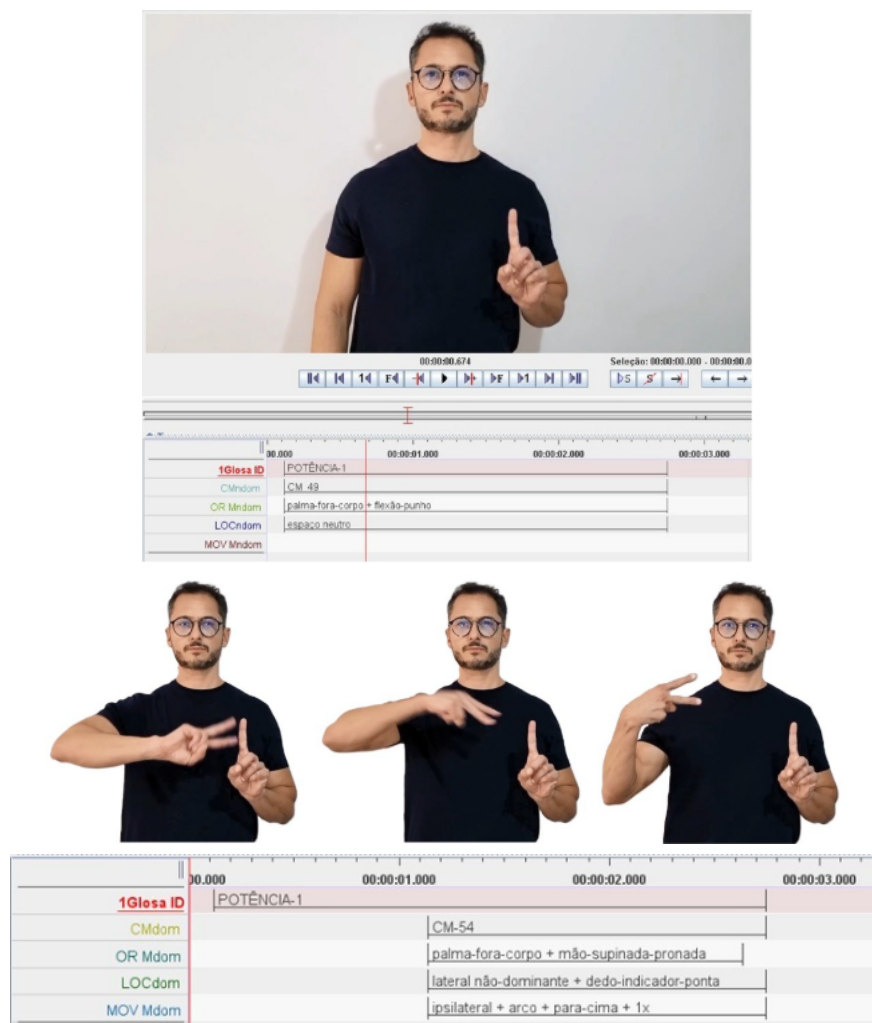


Figura 3: Sequência temporal das imagens capturadas de um sinal-termo para POTÊNCIA

Finalizada a descrição fonético-fonológica no *ELAN* com base nos códigos selecionados, inicia-se a realização da terceira camada direcionada à elaboração de fichas com descrição fonético-fonológica.

Para ilustrar como tais fichas foram elaboradas, escolhemos os termos POTÊNCIA e CUBO. Essa escolha se justificou uma vez que o primeiro permite observar como conceitos mais abstratos podem receber soluções sinalizadas diversas (mais motivadas por convenções escolares, por relações icônicas ou por aproximações com operações matemáticas) e a segunda, por sua vez, evidencia como conceitos com forte dimensão visual e geométrica também apresentam variação (seja na configuração de mão, seja na orientação espacial ou no movimento).

ID-SINAL	GLOSA ID	LAT-Mao	CM		OR		LOC		MOV		
			Dom	Ndom	Dom	Ndom	Dom	Ndom	Dom	Dom	Ndom
PO-0001	POTÊNCIA-1	2a	54 	49 	palma→corpo, supina, prona	palma→corpo, flex		lateral não-dominante, dedo indicador	espaço neutro	ipsi, 1, arco	—
PO-0002	POTÊNCIA-2	1	49 	54 	palma→corpo, supina, prona			espaço frontal, olho		ipsi, 1, arco, repetido	
			54 	78 	palma→corpo, supina, prona			espaço frontal, olho		ipsi, 1, arco, repetido	
			78 		palma→corpo, supina, prona			espaço frontal, olho		ipsi, 1, arco, repetido	
PO-0003	POTÊNCIA-3	2a	54 	05 	palma→corpo, supina, prona		palma→corpo, flex	lateral não-dominante, dedo mínimo	espaço frontal, olho	ipsi, 1, arco	—
PO-0004	POTÊNCIA-4	2a		69 			palma↑, lat_ulin		lateral dominante, peito		contra, toque duplo
			49 	54 	palma→corpo, supina, prona			lateral não-dominante, palma mão traca		ipsi, 1, arco, repetido	
			54 	69 	palma→corpo, supina, prona		palma→corpo	lateral não-dominante, palma mão traca	espaço frontal, pescoço	ipsi, 1, arco, repetido	
			78 	04 	palma→corpo, supina, prona			lateral não-dominante, palma mão traca		ipsi, 1, arco, repetido	
PO-0005	POTÊNCIA-5	2a		69 			palma→corpo		espaço neutro, lateral não-dominante		inicial, → final, motivado, ipsi, 1, arco
PO-0006	POTÊNCIA-6	2a	54 	05 	palma→corpo, supina, prona		palma→corpo	lateral não-dominante, dedo mínimo	espaço neutro, lateral não-dominante	ipsi, 1, arco	—
			54 	49 	palma↓, supina, prona		palma→corpo	lateral não-dominante, dedo mínimo	espaço neutro	ipsi, reto, repetido	—

Figura 4: Exemplo de ficha fonológica dos sinais do termo POTÊNCIA

ID-SINAL	GLOSA ID	LAT-Mao	CM		OR		LOC		MOV		ENM
			Dom	Ndom	Dom	Ndom	Dom	Ndom	Dom	Ndom	
CB-0001	CUBO-1	2s	01 	01 	palma_f2, flex	palma_f2, flex	plano vertical, espaço frontal, ombro, R-loc: médio	plano vertical, espaço frontal, ombro, R-loc: médio	—	—	
					palma_f_dorso, ext	palma→corpo, ext	plano horizontal, espaço neutro, R-loc: médio	plano horizontal, espaço neutro, R-loc: médio	encont. ↓, →espaço, cont none	encont. ↓, →espaço, cont none	
CB-0002	CUBO-2	2s	01 	01 	palma_f2, flex	palma_f2, flex	plano vertical, espaço frontal, ombro, R-loc: médio	plano vertical, espaço frontal, ombro, R-loc: médio	—	—	
					palma_f_dorso, ext	palma↓, ext	espaço frontal, lateral não-dom, R-loc: médio	espaço frontal, lateral domin, R-loc: médio	encont. ↑, arco, cont none	encont. ↓, arco, cont none	
CB-0003	CUBO-3	2s	01 	01 	palma_f2, flex	palma_f2, flex	plano horizontal, espaço neutro, R-loc: médio	plano horizontal, espaço neutro, R-loc: médio	—	—	
					palma_f_dorso, ext	palma→corpo, ext	plano horizontal, espaço neutro, R-loc: médio	plano horizontal, espaço neutro, R-loc: médio	encont. ↓, →espaço, cont none	encont. ↓, →espaço, cont none	
CB-0004	CUBO-4	2s	01 	01 	palma_f2, flex	palma_f2, flex	plano horizontal, espaço neutro, R-loc: médio	plano horizontal, espaço neutro, R-loc: médio	—	—	
			76 	76 	palma_f2, ext	palma_f2, ext	espaço neutro, lateral não-dominante, R-loc: médio	espaço neutro, lateral dominante, R-loc: médio	encont. ↑, arco, cont none	encont. ↓, arco, cont none	
CB-0005	CUBO-5	2s	02 	02 	palma_f2, flex	palma_f2, flex	plano horizontal, espaço neutro, R-loc: médio	plano horizontal, espaço neutro, R-loc: médio	—	—	
			75 	75 	palma_f_dorso, ext	palma→corpo, ext	plano horizontal, espaço neutro, R-loc: próximo	plano horizontal, espaço neutro, R-loc: próximo	encont. →, reto, cont none	encont. reto, →espaço, cont none	
					palma_f2, prona, supina	palma_f2, prona, supina	espaço neutro, lateral não-dominante, R-loc: médio	espaço neutro, lateral dominante, R-loc: médio	encont. inicial, tl, circular, cont none, 2x	encont. inicial, tl, circular, cont none, 2x	

Figura 5: Exemplo de ficha fonológica dos sinais do termo CUBO

As fichas fonológicas demonstram a passagem do registro documental para a descrição analítica. Ao registrar configuração de mão, orientação, locação, movimento e expressões não manuais, o *corpus* torna visíveis diferenças que poderiam desaparecer em uma

tradução por palavra. Essa é uma condição importante para estudos comparativos, para análise da variação e para futuras observações pedagógicas sobre reconhecimento e uso dos sinais por estudantes Surdos.

Do ponto de vista de replicação, uma recomendação seria a de que cada estudo definisse previamente o nível de detalhamento adotado na descrição dos dados. Nem todo *corpus* precisará registrar o mesmo conjunto de subcategorias. Contudo, deve explicitar quais parâmetros serão descritos, quais fontes normativas ou catálogos serão adotados e como divergências de anotação serão resolvidas. Essa explicitação permite que os resultados sejam avaliados, comparados e eventualmente ampliados por outros pesquisadores.

Do ponto de vista linguístico-pedagógico, essa descrição antecipa perguntas que poderiam ser feitas na fase seguinte, a fase de avaliação dos sinais por professores Surdos, tais como: quais traços do sinal favorecem o reconhecimento do conceito? Que variantes são mais transparentes para os sujeitos Surdos? No papel de professores, como eles negociam diferenças entre sinais conhecidos localmente e aqueles presentes em materiais institucionais? A ficha fonológica, portanto, não é apenas instrumento de arquivo, mas, dentro da metodologia descrita funciona para preparar a observação da língua em uso por seus usuários.

3.3 Fase 3: Avaliação por professores Surdos

A terceira fase desloca o *corpus* da documentação e da descrição para a avaliação comunitária. Um *corpus* constituído apenas por levantamento documental poderia reunir sinais de fontes diversas sem saber como esses sinais são reconhecidos, usados ou avaliados por sujeitos que atuam no ensino de matemática em Libras. Por isso, a pesquisa submeteu o repertório organizado à avaliação de professores Surdos.

A fase foi concretizada por meio de questionário bilíngue, em português e Libras, destinado a verificar o reconhecimento de sinais, uso efetivo, adequação ao termo matemático e possíveis alternativas conhecidas pelos participantes. A escolha por professores Surdos tem fundamento metodológico e epistemológico conforme o entendimento de que esses sujeitos possuem experiência profissional com a Libras, com a matemática e com a mediação pedagógica de conceitos especializados. O mais importante nessa fase é entender que não é concebida com o intuito de requerer que os professores “inventem” sinais artificialmente,

como tem ocorrido em alguns projetos de levantamento terminológico atuais. Os participantes não são convocados a criar sinais para suprir questões lacunares definidas pela língua portuguesa ou por demandas institucionais. Em outra direção, a proposta é a de avaliar repertórios documentados, considerando inteligibilidade, recorrência, adequação conceitual, circulação e potencial pedagógico, o que reconhece os professores Surdos como agentes epistêmicos e não como informantes convocados para preencher lacunas lexicais desconsiderando a complexidade linguístico-identitária que a língua carrega e seu funcionamento sociocultural.

Em estudos replicáveis, a validação comunitária pode assumir diferentes formatos, como questionários bilíngues, entrevistas com especialistas Surdos, grupos focais, oficinas de análise ou sessões de julgamento de aceitabilidade. O ponto fundamental é que a etapa de avaliação deve ser planejada de modo coerente com o perfil dos participantes e com o objetivo do *corpus*. Quando o *corpus* tem finalidade pedagógica, participantes com experiência de ensino na área nos parecem ser os interlocutores centrais.

Na pesquisa de doutorado que deriva este desenho metodológico, um questionário foi concebido como instrumento de avaliação comunitária do repertório. Suas perguntas permitiram distinguir níveis de relação com o sinal, de modo que: (a) conhecer não implica necessariamente usar; (b) usar não significa considerar a forma mais adequada; e (c) considerar adequado não pressupõe que a variante tenha ampla circulação. Essa distinção é metodologicamente produtiva porque impede que a avaliação seja reduzida a respostas binárias.

Componente	Tipo de Conteúdo	Descrição do Conteúdo	Acesso
Vídeo de Apresentação	Vídeo (Libras)	Introdução sobre a pesquisa	
Vídeo Explicativo 1	Vídeo (Libras)	Orientação sobre as perguntas	
Vídeo Explicativo 2	Vídeo (Libras)	Instrução sobre as respostas	
Vídeos dos Sinais	Vídeo (Libras)	Apresentação de 64 sinais ao todo	
Opções de Resposta	Checkbox Múltiplo	Três colunas de avaliação por sinal: (A) Conhece; (B) Usa; (C) Valida	
Resposta Aberta	Espaço aberto	Espaço para escrever suas observações	

Quadro 5 - Desenho estrutural do questionário com acesso disponível em QR-code

Essa avaliação não funciona como certificação normativa de um sinal único. Ela acrescenta uma camada de inteligibilidade comunitária ao *corpus*, indicando quais variantes são mais reconhecidas, quais têm baixa circulação, quais são percebidas como mais

adequadas e quais ainda demandam algum nível de discussão. No caso de POTÊNCIA, por exemplo, as respostas¹ evidenciaram baixa convergência entre conhecer, usar e considerar adequado determinado sinal. Tal resultado não é um problema a ser apagado, mas um dado que revela circulação irregular, repertórios locais e necessidade de mediação pedagógica.

A validação comunitária também cria ponte com a fase posterior. Ao avaliar sinais, os professores Surdos não apenas conferem a pertinência de itens lexicais, mas ajudam a identificar repertórios que poderão ser mobilizados, comparados e problematizados em sala de aula ou outros contextos de ensino e uso da língua. Assim, a terceira fase prepara a análise da língua em uso, em vez de encerrar o percurso em uma lista de sinais.

Na pesquisa da qual este recorte foi extraído, a validação por professores Surdos não encerra a discussão sobre os sinais, mas qualifica o repertório para novos usos analíticos. Em vez de produzir uma lista final, a fase 3 oferece uma base interpretativa para compreender a circulação dos sinais e para formular novas questões de pesquisas e outros objetivos investigativos.

Essa ponte é importante para que a quarta fase da pesquisa não apareça como deslocamento abrupto. A título de ilustração, com base na pesquisa em sua integralidade, parcialmente aqui descrita, o que foi observado na fase seguinte dependeu, em parte, da qualidade do repertório previamente documentado. Se a fase 3 indica sinais reconhecidos, sinais pouco usados e sinais controversos, a fase seguinte buscou observar como esses repertórios entram em interação com os sujeitos a partir de tarefas, materiais visuais e mediações realizadas pelo pesquisador para alcançar os objetivos e responder as perguntas que a pesquisa se propôs responder. Outras formas de validação podem ser elaboradas e incluir questões mais específicas.

No estudo referido neste artigo, cujo objetivo foi o de “investigar como sujeitos Surdos constroem sentidos e significados de termos matemáticos em Libras por meio de práticas colaborativas e multimodais, em contextos sociais”, a avaliação ofereceu uma base para o desenvolvimento da quarta fase. Nela, buscou-se analisar como os professores reconhecem, mobilizam, modificam ou negociam esses repertórios em situação concretas de ensino-aprendizagem de matemática a partir de encontros elaborados em forma de práticas situadas com base na pedagogia dos multiletramentos (NLG, 1996; Kalantzis; Cope; Pinheiro, 2020). Outros estudos que escolham usar o modelo aqui apresentado, adaptando os passos

metodológicos descritos, mesmo que com outros objetivos analíticos, terão uma base documental de *corpus* robusta, que poderá levar em conta o respeito à variação e diversidade próprias da cultura e da identidade surda, porém sem desconsiderar o rigor que o documentar de línguas de sinais necessita para respeitar sua natureza visuoespacial e multimodal.

4. Da descrição à aplicação: por que a metodologia importa?

A metodologia apresentada neste artigo não deve ser compreendida como modelo fechado, a ser reproduzido linearmente em qualquer investigação que busque se amparar na constituição de *corpora* em Libras como base para quaisquer objetivos. Seu alcance reside em funcionar como arquitetura metodológica adaptável, capaz de orientar pesquisas que documentem repertórios terminológicos em diferentes áreas, sem reduzir a Libras a equivalências lexicais.

A passagem da descrição à aplicação pedagógica depende da coerência entre objeto, fontes, participantes, parâmetros descritivos e finalidade social da pesquisa. Ao registrar ocorrências, preservar a materialidade visual-espacial dos sinais, descrever parâmetros fonético-fonológicos e submeter variantes à avaliação de professores Surdos, a metodologia cria condições para que estudos futuros observem como esses repertórios podem ser mobilizados, reconhecidos, negociados ou transformados em práticas pedagógicas.

Em vez de apresentar um manual de replicação, apresentamos um protocolo orientado por perguntas. As questões funcionam como pontos de partida para pesquisadores que queiram se valer do aparato metodológico aqui descrito para verificar rastreabilidade, diversidade documental, rigor descritivo, participação surda e abertura à variação, preservando a flexibilidade necessária a cada contexto investigativo.

Fase	Perguntas orientadoras
Fase 1	a) As fontes são diversas (escritas, visuais, videográficas, acadêmicas, pedagógicas, comunitárias, institucionais e digitais)?
	b) É possível identificar autoria, data, suporte, contexto e forma de acesso?
	c) Que condições autorizam a entrada de uma fonte no corpus?
	d) Como serão tratadas fontes instáveis, como vídeos online, redes sociais ou materiais sem autoria explícita?
	a) Cada ocorrência relaciona termo, sinal, fonte, variante, glosa e contexto de uso?
	b) O sistema permite rastrear a origem do dado e comparar variantes de um mesmo termo?

Fase	Perguntas orientadoras
Fase 2	c) Quais parâmetros serão descritos? d) A descrição contempla lateralidade, configuração de mão, orientação, locação, movimento e expressões não manuais? e) Os códigos adotados são consistentes e justificados? f) O percurso do dado pode ser reconstruído da fonte à ficha documental e fonológica? g) As decisões de anotação, dúvidas e ambiguidades ficam registradas?
Fase 3	a) Quem avalia os sinais e que experiências mobiliza? b) A avaliação analisa circulação, inteligibilidade, adequação e uso, em vez de induzir a criação artificial de sinais inexistentes? c) A variação é tratada como dado linguístico ou como problema a eliminar? d) As decisões metodológicas preservam a identidade surda e evitam homogeneização dos repertórios?
Fase(s) Seguinte(s)	a) Como o <i>corpus</i> poderá subsidiar ensino, aprendizagem, tradução, interpretação, formação docente e novas pesquisas científicas que se valem de repertório lexical técnico-científico especializado? b) O que deve ser mantido para garantir rigor e o que pode ser adaptado ao novo domínio?

Quadro 6 – Protocolo com perguntas direcionadoras para a replicação da metodologia

As perguntas reunidas no protocolo deslocam a noção de replicabilidade de uma lógica de reprodução mecânica para uma lógica de adaptação responsável. Replicar, nesse caso, significa preservar princípios de rigor, tais como: diversificar fontes, registrar metadados, manter a rastreabilidade das ocorrências, descrever a materialidade visual-espacial dos sinais, reconhecer a variação, explicitar decisões analíticas e envolver sujeitos Surdos como agentes epistêmicos.

Esse modo de compreender a replicabilidade evita transformar o *corpus* em banco de sinais desvinculado dos contextos de circulação ou converter a metodologia em roteiro fixo. O protocolo por perguntas sustenta critérios de rigor, mas preserva a possibilidade de reformulação conforme os propósitos da pesquisa, os materiais disponíveis e as comunidades envolvidas.

A metodologia importa porque transforma a descrição em possibilidade de uso, não como aplicação automática, mas como base para práticas futuras mais responsivas à língua em uso, à variação e à identidade surda.

5 considerações finais

Este artigo apresentou o desenho metodológico de constituição de um *corpus* de sinais-termo matemáticos em Libras, focalizando as três primeiras fases de uma pesquisa de

doutorado, da qual este trabalho constitui um recorte parcial. A discussão mostrou que levantamento documental, organização do *corpus*, descrição fonético-fonológica e avaliação comunitária não são procedimentos isolados, mas camadas articuladas de um mesmo percurso, cuja lógica pode ser adaptada a outros domínios terminológicos em Libras.

A primeira fase reconhece repertórios em circulação; a segunda transforma ocorrências dispersas em dados rastreáveis e comparáveis; e a terceira submete esse repertório à leitura de professores Surdos. Em conjunto, essas fases constroem uma base que não cria sinais artificialmente, não impõe padronização e não trata a variação como ruído. Ao contrário, documentam a língua viva, preservam sua heterogeneidade e criam condições para futuras análises pedagógicas.

A versão aqui proposta também reforça que a inclusão educacional de estudantes Surdos não pode ser dissociada da manutenção da identidade surda. Incluir, nesse contexto, não significa homogeneizar repertórios para adequá-los a uma norma externa, mas produzir instrumentos que permitam transitar entre variantes, contextos escolares, materiais avaliativos e práticas comunitárias. O *corpus* torna-se, assim, uma ferramenta metodológica, política e pedagógica: documenta para incluir sem apagar.

Por fim, o artigo oferece um caminho replicável para a constituição de *corpora* terminológicos em Libras em outras áreas do conhecimento. Ao propor uma arquitetura baseada em documentação multimodal, rastreabilidade, anotação videográfica, descrição fonético-fonológica e validação comunitária, o estudo contribui para ampliar práticas de pesquisa que respeitem a materialidade visual-espacial da Libras e reconheçam sujeitos Surdos como produtores de conhecimento.

Referências

- ATAYDE, S. T. S. *O uso da Libras na matemática do fundamental: uma proposta de glossário*. 2019. 189 f. Dissertação (Mestrado em Matemática em Rede Nacional) - Universidade Federal de Goiás, Catalão, 2019.
- BERBER SARDINHA, T. *Linguística de Corpus*. Manole, 2004.
- BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics - Investigating language structure and use*. Cambridge: Cambridge University Press, 1998.
- CARVALHO, D. C. T. de. *Calculibras – construindo um glossário de Matemática em Libras na Web*. 2017. 112 f. Dissertação (Mestrado em Diversidade e Inclusão) – Instituto de Biologia, Universidade

Federal Fluminense, Niterói, 2017.

CASTRO JÚNIOR, G. de. *Variação linguística em Língua de Sinais Brasileira: foco no léxico*. 2011. 123 f. Dissertação (Mestrado em Linguística) – Universidade de Brasília, Brasília, 2011.

CASTRO JÚNIOR, G. de. *Projeto Varlibras*. 2014. 259 f. Tese (Doutorado em Linguística) – Universidade de Brasília, Brasília, 2014.

CRASBORN, O.; FIÉVEZ, H. Corpus Approaches to Sign Language Linguistics. In: *Oxford Research Encyclopedia of Linguistics*. Oxford University Press, 2021.

D'AZEVEDO, R. P. *Terminologia da matemática em Língua de Sinais Brasileira: proposta de glossário bilíngue Libras-português*. 2019. 322 f. Dissertação (Mestrado em Linguística) – Instituto de Letras, Universidade de Brasília, Brasília, 2019.

FAULSTICH, E. Especificidades semânticas e lexicais: a criação de sinais-termos na Língua Brasileira de Sinais. In: BIDARRA, J; MARTINS, T. A.; SEIDE, M. S. (orgs) *Entre a Libras e o Português: desafios face ao bilinguismo*. Londrina: EDUEL, 2016. p. 69 – 81.

FAULSTICH, E. *Sinal-Termo*. Nota lexical. Brasília: Centro Lexterm, 2012. Disponível em: <http://www.centrolexterm.com.br/notas-lexicais>. Acesso em: 11 ago. 2026.

FLORES, V. M.; REBECHI, R. R. Linguística de *Corpus* aplicada à Libras: ferramentas de transcrição para criação de corpus. ReVEL, edição especial, v. 21, n. 20, p. 1-25, 2023. Disponível em: <https://lume.ufrgs.br/handle/10183/271464>. Acesso em: 11 ago. 2026.

GEERTZ, C. *The interpretation of cultures: selected essays*. New York: Basic Books, 1973. 476 p.

HAUSER, P. C.; O'HEARN, A.; MCKEE, M.; STEIDER, A.; THEW, D. Deaf epistemology: Deafhood and deafness. *American Annals of the Deaf*, v. 154, n. 5, p. 486-492, 2010.

JEWITT, C. Multimodality and Literacy in School Classrooms. *Review of Research in Education*, v. 32, n. 1, p. 241-267, 2008. Disponível em: <https://journals.sagepub.com/doi/abs/10.3102/0091732x07310586>. Acesso em: 11 ago. 2026.

KALANTZIS, M.; COPE, B.; PINHEIRO, P. *Letramentos*. Campinas: Editora da Unicamp, 2020.

KRESS, G. *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge, 2010.

KRESS, G.; VAN LEEUWEN, T. *Multimodal discourse: the modes and media of contemporary communication*. London: Arnold, 2001.

LADD, P. *Understanding Deaf Culture: In Search of Deafhood*. Multilingual Matters, 2003.

LEITE, T. de A.; QUADROS, R. M. de. Línguas de sinais do Brasil: reflexões sobre o seu estatuto de risco e a importância da documentação. In: *Estudos da Língua de Sinais*. Florianópolis: Editora Insular, v. II, p. 15-27, 2014.

LOBATO, M. J. S. *Educação bilíngue no contexto escolar inclusivo: a construção de um glossário em*

Libras e Língua Portuguesa na área de matemática. 2015. 257f. Dissertação (Mestrado Profissional em Ensino de Ciências Naturais e Matemática) - UFRN, Natal, 2015.

LODI, A. C. B.; HARRISON, K. M.; CAMPOS, S. R. L. de. *Língua de sinais e escola: sociolinguística e políticas linguísticas*. Cadernos de Educação, v. 51, p. 12-29, 2015.

MARTINS, L. A. *Educação matemática para Surdos: contribuições de um glossário para o ensino de probabilidade e estatística*. 2019. 113f. Dissertação (Programa de Pós-Graduação para Ciências e Matemática) – IFG, Jataí, 2019.

MOITA LOPES, L. P. Linguística aplicada e vida contemporânea: problematização dos construtos que tem orientado a pesquisa. In: MOITA LOPES, L. P. *Por uma lingüística aplicada indisciplinar*. São Paulo: Parábola, 2006. p. 85-105.

MOITA LOPES, L. P. Fotografias da Linguística Aplicada brasileira na modernidade recente. In: MOITA LOPES, L. P. *Linguística aplicada na modernidade recente*. São Paulo: Parábola Editorial, 2013. p. 15-38.

NLG. A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, v. 66, n. 1, p. 60-92, 1996.

PADDEN, C.; HUMPHRIES, T. *Deaf in america, voices from a culture*. Cambridge: Harvard University Press, 2000.

PIKE, K. L. *Language in relation to a unified theory of the structure of human behavior*. Mouton, 1967.

QUADROS, R. M. de. *Proposta de manual de transcrição do Corpus Libras: Inventário Nacional*. Florianópolis: UFSC, 2015.

QUADROS, R. M. de. Documentação da Língua Brasileira de Sinais. In: Seminário Ibero-americano de Diversidade Linguística, 2014, Foz do Iguaçu. *Anais [...]*. Brasília: IPHAN, 2016. p. 157-173. Disponível em: <https://dspace.unila.edu.br/items/576640c1-d5b2-43f4-acf1-152724893e5e>. Acesso em: 11 ago. 2026.

QUADROS, R. M. de; KARNOPP, L. B. *Língua de sinais brasileira: estudos linguísticos*. Porto Alegre: Artmed, 2004.

SCHEMBRI, A.; FENLON, J.; RENTELIS, R.; CORMIER, K. Building the British Sign Language Corpus. *Language Documentation & Conservation*, v. 7, p. 136-153, 2013.

STROBEL, K. *As imagens do outro sobre a cultura surda*. Florianópolis: UFSC, 2008.

STUMPF, M. R.; PIZZIO, A. L.; LUCINDA, J. O.; QUADROS, R. M. de; CRASBORN, O. *SignBank da Libras*. *Fórum Linguístico*, Florianópolis, v. 17, n. 4, p. 5475-5487, out/dez. 2020. Disponível em: <https://periodicos.ufsc.br/index.php/forum/article/view/77342>. Acesso em: 11 ago. 2026.

XAVIER, A. N.; BARBOSA, P. A. Diferentes pronúncias em uma língua não sonora? Um estudo da variação na produção de sinais da Libras. *DELTA: Documentação de Estudos em Linguística Teórica e*

Aplicada, São Paulo, v. 30, n. 2, p. 371-413, 2014. Disponível em: <https://revistas.pucsp.br/index.php/delta/article/view/17784>. Acesso em: 11 ago. 2026.

YONAHA, T. Q. A pesquisa documental como ferramenta metodológica na linguística aplicada. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, São Paulo, v. 40, n. 1, 2024. Disponível em: <https://revistas.pucsp.br/index.php/delta/article/view/63694>. Acesso em: 11 ago. 2026.

ABSTRACT: This article presents the methodological design developed in a doctoral research project aimed at building a corpus of mathematical sign-terms in Libras. To this end, it outlines a replicable methodological framework for studies seeking to document terminological corpora in Brazilian Sign Language, bringing together multimodal documentary research, Corpus Linguistics, video annotation, phonetic-phonological description, and community-based evaluation by Deaf teachers. The article argues that building a corpus in Libras should not be understood as the random creation of signs or as an instrument of imposed standardization, but rather as the situated documentation of repertoires already in circulation. By focusing on the first three phases of the research, the article shows how each methodological decision contributes to preserving the visual-spatial materiality of Libras, recognizing linguistic variation, valuing Deaf identity, and laying the groundwork for future inclusive pedagogical practices.

Keywords: Libras corpus; Sign-terms; Methodology; Deaf identity; Inclusion.