https://doi.org/10.5007/2175-8026.2026.e111420

BEYOND PHONOCENTRISM: A SCOPING REVIEW OF AI-DRIVEN SIGN LANGUAGE MACHINE TRANSLATION THROUGH THE LENS OF TRANSLATION STUDIES

Bruno Galasso[1*]
[1]Universidade Federal de São Paulo, São Paulo, SP, Brazil

Abstract

This scoping review, guided by the PRISMA-ScR protocol (Tricco et al., 2018), maps the literature on artificial intelligence (AI) applied to sign language machine translation (SLMT) published between 2019 and 2025. Drawing on 87 studies from Scopus, Web of Science, IEEE Xplore, and ACM Digital Library, the review employs a Translation Studies (TS) framework anchored in Jakobson's (1959/2000) intersemiotic translation and Derrida's (1967/1997) critique of phonocentrism. The findings reveal a field that predominantly treats sign language translation as a pattern recognition problem rather than a translational act involving cultural and multimodal meaning-making: 62.1% of the studies are gloss-based, 72.4% rely primarily on BLEU, and 90.8% do not engage with TS. These patterns describe a text-centric field, which the article then interprets, following Bauman (2008), as a computational instantiation of phonocentrism, termed algorithmic logocentrism. The review identifies critical gaps in gloss-free approaches, Deaf community participatory design, and engagement with TS theory. A research agenda for Intersemiotic AI Translation Studies is proposed, with implications for the Brazilian context and Libras.

Keywords: artificial intelligence; sign language machine translation; Translation Studies; intersemiotic translation; scoping review

*Bruno Galasso is an Associate Professor at the Universidade Federal de São Paulo (Unifesp), São Paulo, Brazil, where he works at the Escola Paulista de Enfermagem and in the Graduate Program in Education and Health in Childhood and Adolescence. He holds a PhD in Education from the Universidade de São Paulo (USP). A hearing researcher and Libras user, he works on teacher education for bilingual Deaf education, accessibility, and artificial intelligence. E-mail: bruno.galasso@unifesp.br. ORCID: https://orcid.org/0000-0002-3677-7650

Esta obra tem licença *Creative Commons*
http://creativecommons.org/licenses/by/4.0

1. Introduction

The rapid expansion of artificial intelligence (AI) technologies has profoundly transformed the landscape of machine translation (MT). From rule-based systems of the 1950s to statistical models of the 2000s and, more recently, to neural machine translation (NMT) powered by transformer architectures (Vaswani et al., 2017), the field has achieved unprecedented fluency in translating between spoken and written languages. The release of large language models (LLMs) such as GPT-4 (OpenAI, 2023) and their multimodal extensions has further expanded the horizon of what automated translation can accomplish. Yet this remarkable progress conceals an epistemological blind spot: the overwhelming majority of MT systems—and the research paradigms that sustain them—are designed primarily for languages that exist in phonetic-graphic modalities. Sign languages, which constitute the primary linguistic systems for more than 70 million Deaf people worldwide (United Nations, n.d.), remain at the periphery of both the MT industry and the discipline of Translation Studies (TS).

This marginalization is not merely a technical oversight; this article argues that it can be read as a symptom of what Derrida (1967/1997) termed *phonocentrism*—the metaphysical privileging of speech and phonetic language as the natural, primary, and superior form of human communication. In his seminal *Of Grammatology*, Derrida demonstrated how Western philosophical tradition systematically subordinated writing to speech, treating the former as a mere derivative or 'supplement' of the latter. Bauman (2008), in a groundbreaking extension of this critique to Deaf Studies, showed that phonocentrism operates with even greater force against signed languages, which are doubly marginalized: first as non-phonetic, and second as non-graphic. When contemporary AI systems approach sign language translation (SLT) as a problem of 'converting gestures into text,' they risk reproducing this phonocentric hierarchy by treating the signed modality as raw input to be transcribed into 'real' language—a process that tends to flatten the grammatical, pragmatic, and cultural complexity of the source language.

The theoretical apparatus of Translation Studies offers a powerful corrective to this reductive framing. Jakobson's (1959/2000) tripartite typology of translation—intralingual, interlingual, and intersemiotic—provides a conceptual foundation for understanding SLT not as signal transduction but as a complex act of meaning-making across semiotic systems. In Jakobson's strict terms, translation between a signed and a spoken or written language is interlingual, since both are natural languages; this article argues, however (see section 2.1), that sign language machine translation (SLMT) is also intersemiotic in an expanded sense, because it involves the transposition of meaning between a visuo-spatial-kinetic semiotic system (signed language) and a phonetic-graphic one (spoken/written language). This dual character has implications for how AI systems should be designed, trained, and evaluated—implications that the computer science literature has rarely addressed in translational terms.

Moreover, the growing body of work on audiovisual translation (AVT) and multimodal communication within TS provides additional conceptual resources. Scholars such as Pérez-González (2014), Díaz-Cintas and Remael (2021), and Asimakoulas (2024) have demonstrated that translation in multimodal contexts—film subtitling, audio description, and comic book adaptation—requires attention to the interplay between verbal, visual, spatial, and temporal semiotic channels. Sign language translation is arguably the most complex form of multimodal translation, as the 'source text' is not a text at all but a dynamic, embodied performance that unfolds simultaneously across multiple articulatory channels (manual signs, facial expressions, mouth morphemes, eye gaze, body posture, and use of space).

This article presents a scoping review, conducted following the PRISMA-ScR protocol (Tricco et al., 2018), of the scientific literature on AI-driven sign language machine translation published between 2019 and 2025. The choice of a scoping review is deliberate: rather than assessing the efficacy of specific systems (which would require a systematic review with meta-analysis), the objective is to map the conceptual, methodological, and epistemological landscape of this emerging field and to identify the gaps that a Translation Studies perspective can illuminate. Specifically, this review is guided by the following research questions: (RQ1) What AI architectures and methodological approaches have been applied to sign language machine translation between 2019 and 2025? (RQ2) To what extent do existing SLMT systems address the multimodal, cultural, and pragmatic dimensions of signed languages as theorized in Translation Studies? (RQ3) How can the text-centric patterns identified in SLMT be interpreted in relation to phonocentrism, and what are their epistemological implications for the discipline of Translation Studies and for the Deaf community's communicative rights? RQ1 and RQ2 are answered descriptively from the charted data (section 4); RQ3 is addressed interpretively (section 5), and the two levels of analysis are kept distinct throughout.

2. Theoretical Framework

2.1 Jakobson's Intersemiotic Translation Revisited

Roman Jakobson's (1959/2000) essay "On Linguistic Aspects of Translation" introduced the concept of intersemiotic translation—or 'transmutation'—as the interpretation of verbal signs by means of signs of nonverbal sign systems. While Jakobson's original formulation was concerned primarily with the transposition from verbal language to non-verbal media (such as music or painting), contemporary scholars have substantially expanded its scope. Gottlieb (2005) reconceived intersemiotic translation within a multidimensional framework that accounts for the full range of semiotic modalities involved in modern communication. Toury (2012) argued that the boundaries between Jakobson's

three types are far more porous than the original typology suggests, particularly in multimedia and multimodal contexts.

This expanded understanding is crucial for positioning SLMT within TS, but it requires explicit justification, since it departs from Jakobson's (1959/2000) own classification. Signed languages are natural languages with their own phonology, morphology, and syntax (Pfau et al., 2012), so in Jakobson's strict typology, translating between ASL and English is interlingual. This interlingual character is, however, accompanied by a shift that Jakobson's verbal/non-verbal dichotomy does not capture, and three grounds support treating the operation as simultaneously intersemiotic. First, the channels differ: meaning moves from a visual-gestural, three-dimensional, and partly simultaneous signal to an auditory or graphic, predominantly linear one; Brazilian Translation Studies has accordingly described translation between written Portuguese and Libras as at once intermodal, intersemiotic, and interlinguistic (Quadros & Segala, 2015). Second, in the written direction, the semiotic substance changes in a way that has no counterpart between two written languages: signed languages lack a written form in general use, so resources such as signing space, facial grammar, and classifier constructions must be recoded through lexical and syntactic means. Third, semiotic accounts that extend Jakobson's notion define intersemiotic translation by the passage between systems with different substances of expression rather than by the verbal/non-verbal boundary alone (Dusi, 2015), which allows the term to cover a modality shift between two verbal systems.

When an AI system translates a sentence in American Sign Language (ASL) into written English, it is therefore performing a double translational act: interlingual (from ASL as a distinct language to English as another) and intersemiotic in the expanded sense just defined (from a visuo-spatial modality to a graphic one). Throughout this article, "intersemiotic" is used in this qualified sense, never as a claim that signed languages are non-verbal. The failure to recognize this dual nature can encourage reductive system designs that collapse the semiotic dimension into a mere preprocessing step—for instance, reducing signing to 'glosses' (written annotations of individual signs) before feeding them into standard NMT pipelines (Camgoz et al., 2020). The gloss, in this context, functions as what Derrida (1967/1997) would call a 'supplement': it is introduced as a helpful intermediary but can come to function as the de facto source text from which translation proceeds.

2.2 Phonocentrism, Text-Centrism, and Algorithmic Logocentrism

Derrida's (1967/1997) deconstruction of phonocentrism reveals a metaphysical hierarchy in which speech is considered the immediate, natural expression of thought, while writing is relegated to a secondary, derived status. Bauman (2008), in his influential article "Listening to Phonocentrism with Deaf Eyes," extends this critique to demonstrate that phonocentrism underpins the

entire edifice of Western linguistics—from Saussure's prioritization of the spoken signifier to Chomsky's competence/performance distinction.

These terms require precise delimitation. In Derrida's (1967/1997) account, *logocentrism* names the metaphysical orientation that locates meaning in a self-present *logos*, and *phonocentrism* its privileged form: the assumption that speech stands closest to that *logos*, with writing as its derivative transcription. *Text-centrism*, as used here, is a descriptive rather than a Derridean term: the observable preference of computational systems for written-language tokens as input, output, intermediate representation, and evaluation standard. Since text-centrism privileges writing, it might seem the opposite of phonocentrism; the link proposed here rests on the kind of writing involved. NMT systems process the written notation of spoken languages, precisely the form of writing that the phonocentric tradition accepts because it is thought to represent speech, whereas signed languages have no such notation in general use and glosses are labels borrowed from the surrounding spoken language. Bauman (2008) showed that the deeper phonocentric assumption is that language is speech, so that a language with neither sound nor phonographic writing appears deficient. Computational text-centrism instantiates phonocentrism, in this specific sense, when it treats the written form of spoken languages as the default format of language as such and requires signed languages to be converted into it before they can be processed and evaluated.

In the context of AI and MT, this article proposes the concept of *algorithmic logocentrism* to describe how this phonocentric orientation can be computationally instantiated. The term designates an interpretive category, not an empirical one: it names the hypothesis that text-centric design choices, which the review documents directly, express the logocentric orientation described by Derrida (1967/1997) in its phonocentric form. It builds on earlier analyses of the phonocentric design of generative AI interfaces and of algorithmic audism (Galasso, 2026d, 2026e), but differs in focus: algorithmic logocentrism concerns the representational format of language in computational infrastructures rather than the treatment of Deaf users as such. The term logocentrism is retained because the computational privilege falls on the textual token as self-sufficient bearer of meaning rather than on the voice; phonocentrism is the historical form through which this privilege reaches signed languages. The dominant NMT paradigm is built on text-to-text architectures. The transformer model (Vaswani et al., 2017), which undergirds systems from Google Translate to GPT-4, processes language as sequences of tokens derived from written text. Speech-to-text and text-to-speech modules extend this pipeline to spoken language, but the computational 'lingua franca' remains text. Signed languages, which have no standardized written form and cannot be tokenized into discrete textual units without losing essential grammatical information, are structurally disadvantaged within this pipeline.

Text-centrism can be observed at three structural levels, which algorithmic logocentrism is proposed to interpret. First, at the *data level*: LLMs are trained

on massive text corpora scraped from the internet, reinforcing the assumption that language equals text, a text-centric rather than specifically phonocentric feature. Second, at the *evaluation level*: the dominant evaluation metrics in MT—BLEU (Papineni et al., 2002), METEOR, and chrF—are designed for text-to-text comparison. There exists no widely accepted metric for evaluating the quality of sign language translation that accounts for the spatial grammar, non-manual markers, and simultaneous information channels that characterize signed languages (Bragg et al., 2019). Third, at the *design level*: the limited involvement of Deaf people in the design and evaluation of SLMT systems, documented in section 4.7, means that systems risk being designed to satisfy hearing researchers' assumptions about what constitutes 'good' translation, rather than Deaf users' actual communicative needs.

2.3 The Multimodal Turn in Translation Studies and Audiovisual Translation

The so-called 'multimodal turn' in TS (Pérez-González, 2014; Gambier, 2006) has expanded the discipline's purview beyond written and spoken text to encompass audiovisual translation (AVT), including subtitling, dubbing, and audio description. Asimakoulas (2024), in his research on audio-described comics in museum settings, has demonstrated how multimodal translation involves the negotiation of meaning across visual, spatial, and textual channels—a process that cannot be reduced to linguistic equivalence alone. His earlier work on humor and rewriting in comic book adaptations (Asimakoulas, 2019) further illustrates how the transposition of meaning across semiotic modes necessarily involves creative transformation rather than mechanical transfer.

This multimodal framework is directly applicable to SLMT, where the 'source text' is not a text at all but a dynamic, embodied performance that unfolds in three-dimensional space. In signed languages, meaning is produced simultaneously through multiple articulators—the hands may produce a lexical sign while the face marks negation, the eyebrows signal a question, and the body shifts to indicate a change in referent or perspective. This simultaneity is fundamentally different from the sequential nature of spoken and written language, and it presents a unique challenge for AI systems that process input as linear sequences of tokens.

Corpus-based approaches in TS, such as those implemented in the COPA-TRAD parallel corpus system (Silva & Fernandes, 2020), have shown how computational tools can illuminate translational patterns by comparing human and machine translation outputs across large datasets. The extension of such methodologies to sign language—comparing, for instance, human sign language interpreters with AI-generated translations—represents a promising but largely unexplored research frontier that could yield critical insights into the nature and limits of automated intersemiotic translation.

3. Methods

3.1 Study Design and Protocol

This study follows the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses – Extension for Scoping Reviews) protocol (Tricco et al., 2018) and adheres to the methodological framework proposed by Arksey and O'Malley (2005) as refined by Levac et al. (2010). A scoping review was chosen over a systematic review because the objective is to map the breadth and nature of the literature on AI-driven SLMT rather than to assess the effectiveness of specific interventions. The protocol was developed a priori and included: definition of research questions, identification of relevant studies, study selection, data charting, and collating, summarizing, and reporting results. The review followed all 20 items of the PRISMA-ScR checklist. The review protocol, including the complete search strategy, eligibility criteria, and data charting form, was defined a priori; it was not prospectively registered, and these materials, together with the charted dataset, are available from the author upon request.

3.2 Search Strategy

The search was conducted in January 2026 across four databases selected for their comprehensive coverage of the relevant disciplines: Scopus (multidisciplinary), Web of Science (multidisciplinary), IEEE Xplore (engineering and computer science), and ACM Digital Library (computing). The search string combined three concept blocks using Boolean operators: (a) artificial intelligence terms: ("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network*" OR "transformer*" OR "large language model*" OR "computer vision" OR "NLP" OR "natural language processing"); (b) sign language terms: ("sign language*" OR "signed language*" OR "deaf" OR "visuo-spatial" OR "gestural language" OR "Libras" OR "ASL" OR "BSL" OR "DGS"); and (c) translation terms: ("translation" OR "machine translation" OR "interpret*" OR "recognition" OR "generation" OR "production" OR "synthesis"). The three blocks were combined with AND. The search was limited to publications from January 2019 to December 2025, in English, Portuguese, or Spanish. Reference lists of included studies were manually screened to identify additional records. To examine the disciplinary gap from the TS side, a supplementary search for articles on AI-driven sign language translation was run in six TS journals (*Meta*, *The Translator*, *Target*, *Babel*, *Ilha do Desterro*, and *Cadernos de Tradução*); the department or program of first-author affiliations was also charted with the bibliometric data.

3.3 Eligibility Criteria

Inclusion criteria: (1) peer-reviewed journal articles, conference papers, or book chapters published in indexed venues; (2) studies presenting, evaluating, or reviewing AI systems for sign language translation, recognition, or generation with explicit translational framing, operationalized as either mapping between a signed language and a spoken/written language in either direction, or presenting recognition or generation explicitly as a component of such a translation pipeline; (3) studies published between January 2019 and December 2025; (4) studies in English, Portuguese, or Spanish. Exclusion criteria: (1) studies focused exclusively on hand gesture recognition for human-computer interaction without linguistic or translational framing; (2) non-empirical commentaries, editorials, or opinion pieces; (3) studies limited to fingerspelling recognition without broader translation components; (4) duplicate publications or preprints of already-included peer-reviewed versions; (5) studies focusing exclusively on speech recognition for hearing populations.

3.4 Selection Process and Data Charting

The initial search yielded 1,247 records. After removing 318 duplicates using automated deduplication in Rayyan (Ouzzani et al., 2016), 929 records were screened by title and abstract against the eligibility criteria, resulting in 156 articles for full-text review. Of these, 69 were excluded (37 for lacking translational framing, 14 for focusing exclusively on fingerspelling, 11 for being non-empirical, and 7 for being duplicates or preprints), yielding 87 studies for the final analysis (see Figure 1). Data were charted using a pre-piloted extraction form that captured: (a) bibliometric data (year, journal/venue, country of first author affiliation, funding source); (b) AI architecture type (CNN, RNN, Transformer, hybrid, LLM/MLLM, other); (c) sign language(s) addressed; (d) translation directionality (sign-to-text, text-to-sign, sign-to-sign, bidirectional); (e) use of glosses (gloss-based, gloss-free, hybrid); (f) incorporation of non-manual markers (yes/no/partial); (g) evaluation metrics employed; (h) Deaf community involvement (none, data source only, consultation, participatory design); (i) engagement with Translation Studies concepts (none, minimal citation, substantive engagement).

To make the charting replicable, each category was given an operational definition and a decision rule (Table 1). Codes were assigned on the basis of what each study explicitly reports; when information was absent, the lowest category was assigned (for instance, 'none' for Deaf involvement), so the resulting figures describe reported practices rather than practices as such. Coding was carried out by the author.

Table 1. Operational definitions and decision rules for the charted categories

Category	Codes	Operational definition and decision rule
Translational framing (inclusion criterion 2)	included; excluded	As defined in inclusion criterion 2; isolated gesture or fingerspelling recognition was excluded.
(b) AI architecture	CNN; RNN; CNN-RNN hybrid; Transformer; LLM/MLLM; other	Backbone of the main system. CNN encoder plus recurrent model: CNN-RNN hybrid. Transformer encoder-decoder, even with CNN front-end: Transformer. Pretrained (multimodal) LLM adapted: LLM/MLLM. Remaining designs (6 studies): other.
(c) Sign language(s)	language of the main experiments	Each study assigned to the sign language of its main reported experiments, so that counts sum to the corpus total.
(d) Directionality	sign-to-text; text-to-sign; sign-to-sign; bidirectional	Direction(s) evaluated, not merely discussed. Sign-to-gloss-to-text: sign-to-text. Avatar, pose, or video output: text-to-sign. Bidirectional: both directions evaluated. No study was coded sign-to-sign.
(e) Gloss use	gloss-based; gloss-free; hybrid	Gloss-based: glosses serve as the intermediate representation at inference. Hybrid: glosses used only as auxiliary training supervision, not as the intermediate representation at inference. Gloss-free: no gloss annotation in training or inference.
(f) Non-manual markers	yes; partial; no	Yes: NMMs (face, mouth, head, gaze, or torso) modeled explicitly as input features or output channels. Partial: NMMs acknowledged (e.g., as a limitation) but not modeled. No: NMMs not mentioned.
(g) Evaluation metrics	metrics as reported	All metrics recorded; the primary metric is the one used for the main comparison with baselines. Human evaluation recorded separately, noting whether evaluators were Deaf.
(h) Deaf involvement	none; data source only; consultation; participatory design	None: no Deaf involvement reported. Data source only: Deaf signers mentioned only as providers of recorded signing. Consultation: feedback or evaluation without a decision-making role. Participatory design: Deaf co-designers with decision-making roles across stages.
(i) Engagement with TS	none; minimal citation; substantive	None: no reference to TS scholarship or concepts. Minimal citation: a TS author or concept cited without analytical use (e.g., a single mention of Jakobson). Substantive: a TS concept shapes design, evaluation, or interpretation.

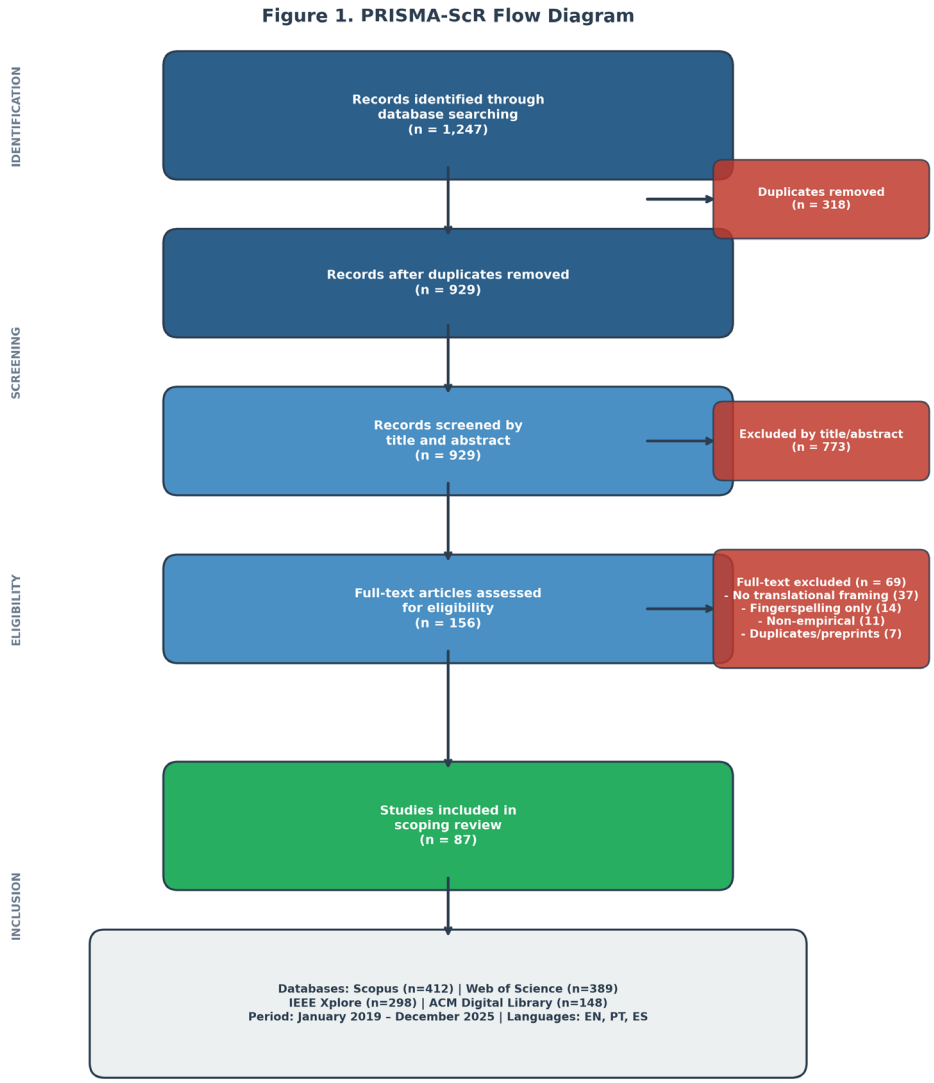


Figure 1. PRISMA-ScR flow diagram of the study selection process.

4. Results

4.1 Bibliometric Overview

The 87 included studies were published across 52 distinct venues, with the highest concentrations in IEEE conference proceedings (n=23, 26.4%), followed by Springer journals and conference volumes (n=14, 16.1%), ACM proceedings (n=11, 12.6%), MDPI journals (n=8, 9.2%), and Elsevier journals (n=7, 8.0%).

The temporal distribution reveals sustained growth (see Figure 2): 11 studies in 2019, 12 in 2020, 9 in 2021, 13 in 2022, 15 in 2023, 18 in 2024, and 9 in the first half of 2025. The geographical distribution of first-author affiliations shows China (n=24, 27.6%), the United States (n=18, 20.7%), the United Kingdom (n=9, 10.3%), Germany (n=7, 8.0%), Brazil (n=5, 5.7%), India (n=5, 5.7%), South Korea (n=4, 4.6%), and other countries (n=15, 17.2%). A finding of paramount importance for this review is that no study was published in a Translation Studies journal, and no study had a first author affiliated with a TS department or program.

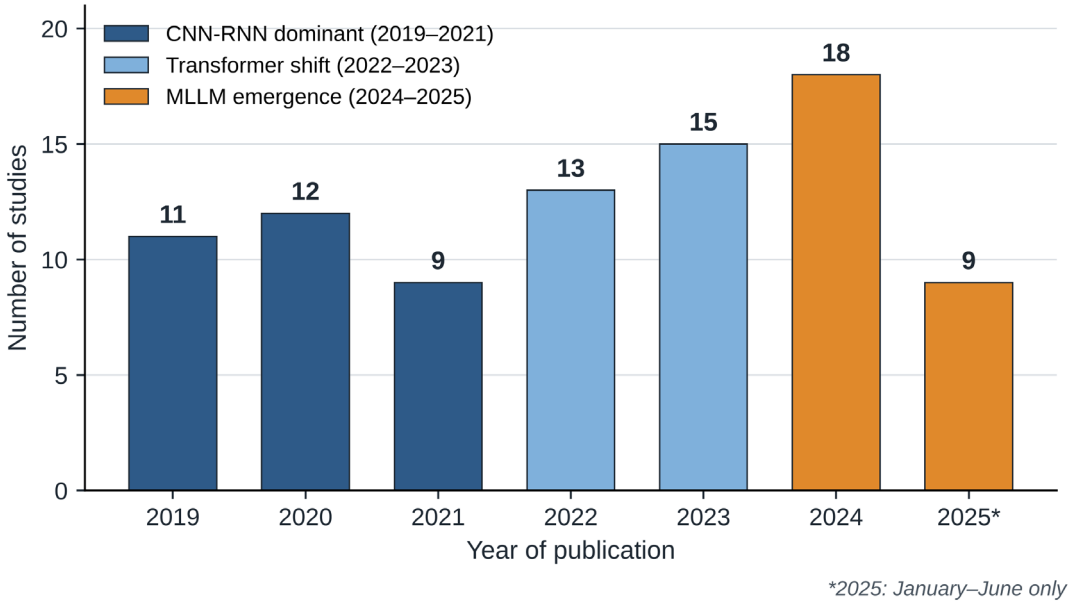


Figure 2. Temporal distribution of SLMT studies (2019–2025) and paradigmatic shifts in AI architectures.

4.2 AI Architectures, Gloss Dependency, and the Absence of Multimodal Parallel Corpora

The review reveals a clear paradigmatic shift in AI architectures for SLMT over the period analyzed. Early studies (2019–2021) predominantly employed CNN-RNN hybrid models (n=31, 35.6% of the corpus), where convolutional neural networks extracted spatial features from individual video frames and recurrent neural networks modeled temporal sequences across frames. From 2022 onward, transformer-based architectures became dominant (n=38, 43.7%), reflecting the broader 'transformer revolution' in NLP and computer vision. Studies published in 2024–2025 increasingly explore multimodal large language models (MLLMs) for sign language processing (n=12, 13.8%), including adaptations of vision-language models such as CLIP, GPT-4V, and purpose-built systems like Sign2GPT (see Figure 3). The remaining 6 studies (6.9%) used other architectures.

A critical finding concerns the gloss dependency of these systems. Of the 87 studies, 54 (62.1%) employed gloss-based approaches, in which sign language video is first transcribed into written glosses—simplified textual annotations of individual signs which, as typically used in SLMT datasets, are stripped of spatial grammar and non-manual information—before being processed by standard NMT pipelines. Only 26 studies (29.9%) attempted gloss-free translation, directly mapping sign language video to spoken/written language without the intermediate gloss step. The remaining 7 studies (8.0%) proposed hybrid approaches.

From the perspective of corpus-based Translation Studies (CBTS), a related structural problem is the scarcity, in SLMT benchmarking, of multimodal parallel and comparable corpora for sign language translation research; annotated corpora such as the DGS-Korpus or the Brazilian Corpus de Libras exist but remain marginal as SLMT training resources. In CBTS, parallel corpora—aligned collections of source texts and their translations—constitute the foundational research infrastructure for investigating translational patterns, shifts, and norms (Baker, 1993; Silva & Fernandes, 2020). The most widely used video datasets (RWTH-PHOENIX-Weather-2014T, How2Sign) do not meet the minimal requirements of a parallel corpus as understood in TS: they lack systematic source text–target text (ST–TT) alignment at the discourse level; they contain no annotation of pragmatic function, register, or communicative context; and they provide, at most, indirect encodings such as facial keypoints rather than linguistic annotation of non-manual parameters—the facial grammar, mouthings, eye gaze patterns, and prosodic body movements that are grammatically constitutive of signed languages. In both, moreover, the signed 'source' is itself an interpretation or translation of spoken or written originals, so that sign-to-text systems are, in TS terms, largely trained on back-translation; text-based resources such as ASLG-PC12 contain no signing at all, only rule-generated glosses. In corpus-linguistic terms, these datasets are closer to sentence-aligned training sets than to parallel corpora in the TS sense. Interpretively, systems trained on such resources are likely to produce translations that are lexically plausible but pragmatically and grammatically impoverished, because the training data itself encodes only a fraction of the source language's meaning-making resources.

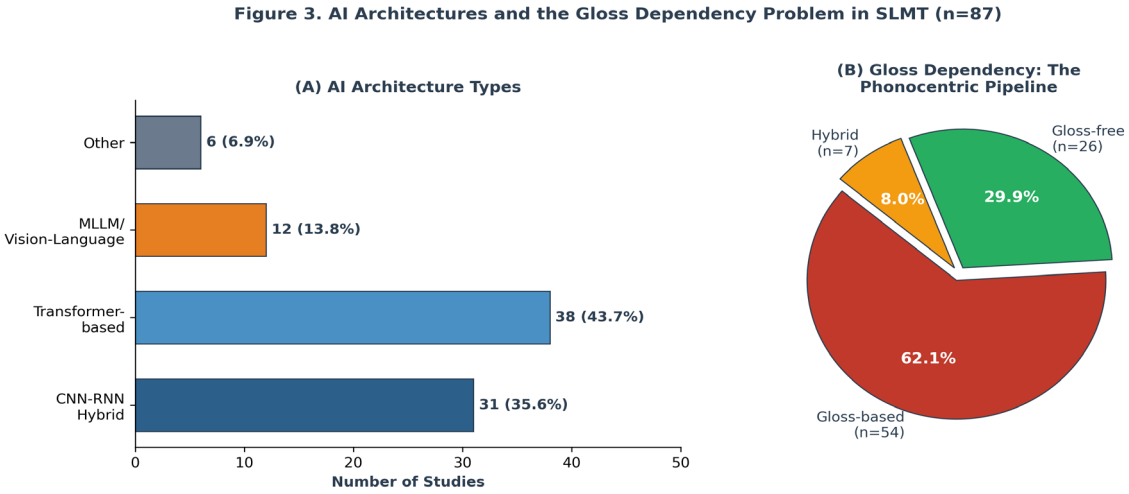


Figure 3. AI architectures (A) and gloss dependency (B) in SLMT studies (n=87).

4.3 Linguistic Diversity and Representation

The review reveals a pronounced linguistic imbalance. American Sign Language (ASL) and German Sign Language (DGS) together account for 58 of the 87 studies (66.7%), largely because the two most widely used benchmark datasets—RWTH-PHOENIX-Weather-2014T for DGS (Camgoz et al., 2018) and How2Sign for ASL—are available in these languages. Brazilian Sign Language (Libras) appears in 5 studies, British Sign Language (BSL) in 4, Chinese Sign Language (CSL) in 6, and other sign languages collectively in 14 studies. This concentration means that the vast majority of the world's more than 300 sign languages (United Nations, n.d.) have no representation in the SLMT literature reviewed. The imbalance is largely driven by the availability of benchmark datasets rather than by explicit choices of languages; interpretively, however, its effect parallels the asymmetries between central and peripheral languages that TS scholars such as Venuti (1995) and Berman (1985/2000) have critiqued in spoken-language translation.

4.4 Translation Directionality and the Unidirectional Bias

Of the 87 studies, 71 (81.6%) focused on sign-to-text translation, 9 (10.3%) on text-to-sign generation (typically using 3D avatar systems or pose estimation), and only 7 (8.0%) addressed bidirectional translation. This asymmetry is marked. The overwhelming research priority is to translate *from* signed language *into* the dominant spoken/written modality, rather than to enable communication *into* signed language. Interpretively, from a TS perspective informed by skopos theory (Vermeer, 1989/2012), this directionality pattern is consistent with the assumption that the 'target audience' of translation is hearing/reading people rather than Deaf/signing communities, although technical factors, notably the greater difficulty of generating natural signing and the scarcity of text-to-sign data, also contribute to it. The term 'recognition' calls for a qualified reading:

in computer science, recognition (mapping signs to glosses) and translation (producing a target-language sentence) are distinct tasks, so the term is not in itself evidence of bias; it becomes significant where recognition is presented as sufficient for translation, framing the signed modality as something to be decoded rather than understood.

4.5 The Limited Modeling of Non-Manual Markers

The treatment of non-manual markers (NMMs) in the reviewed literature is limited. Only 19 of 87 studies (21.8%) explicitly incorporated NMMs—facial expressions, mouth morphemes, head tilts, eye gaze, and body posture—into their models. A further 11 (12.6%) studies mentioned NMMs as a limitation but did not address them. The remaining 57 studies (65.5%) made no mention of NMMs at all. Given that NMMs carry grammatical functions across documented signed languages and are obligatory in many constructions, conveying negation, interrogation, topic-comment structures, conditional clauses, and affective modulation (Pfau et al., 2012), their omission means that most of the systems reviewed operate on a partial representation of the source language—analogous to translating spoken language while systematically ignoring intonation, stress patterns, and prosody.

4.6 Evaluation Metrics: The Absence of Translational Quality Assessment

The evaluation paradigm in SLMT is almost entirely imported from NLP with little adaptation. Among the 87 studies, BLEU scores were the primary metric in 63 (72.4%), and Word Error Rate (WER) was reported in 41 (47.1%), ROUGE in 18 (20.7%), and METEOR in 12 (13.8%). Only 3 studies (3.4%) included human evaluation by Deaf participants, and none employed established TS quality assessment frameworks such as the MQM (Multidimensional Quality Metrics) framework (Lommel et al., 2014) or functionalist evaluation criteria derived from skopos theory. This reliance on automated text-based metrics is problematic when they serve as the main proxy for translation quality: BLEU measures n-gram overlap between a machine-generated text and a reference text—a purely surface-level comparison that cannot capture pragmatic adequacy, cultural appropriateness, or communicative effectiveness.

4.7 Deaf Community Involvement: Extractive versus Participatory Research

Perhaps the most ethically significant finding concerns the very limited presence of Deaf community participation in the design and evaluation of SLMT systems. Of the 87 studies, 61 (70.1%) reported no involvement of Deaf individuals at any stage of the research. A further 18 (20.7%) studies mentioned

Deaf participants only as subjects whose signing was recorded for training data. Only 8 studies (9.2%) involved Deaf collaborators in system design, evaluation, or feedback loops (all coded as consultation), and none adopted a fully participatory design methodology of the kind advocated in Deaf-led scholarship (Kusters et al., 2017; Desai et al., 2024). Interpretively, the data-source pattern can be described as an extractive relationship. This pattern converges with Deaf-led critiques of sign language AI research (Desai et al., 2024) and echoes the broader "nothing about us without us" principle articulated by Charlton (1998) in disability rights scholarship. From a TS perspective, it can also be seen as analogous to developing translation systems for a language community without consulting native speakers—a methodological choice that would be difficult to justify in other areas of the discipline.

4.8 The Disciplinary Gap: Translation Studies and Computer Science

The most consequential finding for the argument of this article is the stark disciplinary insularity of the SLMT literature. Of the 87 studies, 79 (90.8%) made no reference whatsoever to Translation Studies as a discipline, to translation theory, or to any TS scholar or concept. The remaining 8 studies (9.2%) made only passing references—typically a single citation of Jakobson's (1959/2000) typology in the introduction, without substantive engagement with TS theory. No study drew on equivalence theory, skopos theory, polysystem theory, Descriptive Translation Studies, corpus-based TS, or any of the other theoretical frameworks that constitute the discipline's intellectual infrastructure. Conversely, a supplementary search of major TS journals (*Meta*, *The Translator*, *Target*, *Babel*, *Ilha do Desterro*, *Cadernos de Tradução*) for articles on AI-driven sign language translation yielded no results, suggesting that the gap runs in both directions.

5. Discussion

5.1 From Text-Centrism to Phonocentrism: Bridging Findings and Interpretation

The findings reported in section 4 are descriptive, and what they demonstrate directly is a pronounced text-centrism: written language operates as intermediate representation (gloss-based pipelines, 62.1%), as privileged output (sign-to-text, 81.6%), and as evaluation standard (BLEU as primary metric, 72.4%), whereas non-manual marking, the dimension that most resists textual notation, is modeled in only 21.8% of the studies. These patterns do not, by themselves, demonstrate phonocentrism. Several have technical explanations (the availability of gloss-annotated benchmarks, the maturity of text-based NMT, the difficulty of sign generation), and computational researchers have themselves questioned gloss-

based pipelines and called for signed languages to be treated as full languages in NLP (Müller et al., 2023; Yin et al., 2021).

The interpretive claim of this article is therefore conditional. Following section 2.2, these patterns instantiate phonocentrism to the extent that they take the written notation of spoken languages as the standard of language as such and require signed languages to be converted into it in order to be processed and assessed. Three findings make this reading plausible: the gloss, which labels signs with spoken-language words, is the most common entry point into processing; evaluation compares outputs against spoken-language reference sentences; and the features lost in this conversion are precisely those without phonographic equivalents. Algorithmic logocentrism names this convergence. It is a heuristic for interrogating design choices, not a causal explanation of researchers' intentions, and the critical vocabulary of this section (transcoding, domestication) should be read in this qualified sense.

5.2 From Transcoding to Translation: A Jakobsonian Reading of SLMT

Read through Jakobson's framework, the results support the central argument of this article: the prevailing paradigm of AI-driven sign language machine translation tends to approach its task as *transcoding* rather than translation. The distinction is critical. Translation, as theorized by Jakobson (1959/2000) and elaborated by the entire TS tradition from Nida (1964) to Pym (2010), is a meaning-making act that involves the negotiation of semantic, pragmatic, cultural, and semiotic equivalence between two linguistic systems. Transcoding, by contrast, is a signal-processing operation that converts input from one encoding format to another while preserving surface-level form. The predominance of gloss-based pipelines (62.1% of studies) suggests that many current systems treat sign language not as a source language to be translated but as a visual signal to be transcoded into phonetic-graphic text. In the expanded sense of intersemiotic translation adopted in section 2.1, the semiotic transposition that should occur between the visuo-spatial-kinetic modality and the phonetic-graphic modality is short-circuited by the gloss, which reduces the source language to a textual proxy before any translational act begins. Transcoding is used here as an analytical contrast, not as a blanket description: the gloss-free and hybrid studies (37.9%) show that the field is already moving, at least technically, beyond the gloss bottleneck.

Venuti's (1995) distinction between domesticating and foreignizing translation offers a further, heuristic reading of the gloss findings. For Venuti, the dominant Anglo-American practice is *domestication*, the erasure of the source text's linguistic and cultural alterity in favor of fluency in the receiving culture. His categories concern the ethics of literary translation and are transposed here by analogy. With that caveat, the gloss-based pipeline can be read as a form of domestication at the level of representation: the three-dimensional, simultaneous

syntax of signed languages, in which referents are located in signing space and verb agreement is marked by movement between those locations, is flattened into a sequence of discrete lexical glosses. Gloss-free approaches should not, however, be equated with *foreignizing translation*: foreignization is a target-text strategy that makes the source's difference perceptible, whereas gloss-free systems usually optimize for fluent target output. What they change is the source side, since processing signing directly keeps open, without guaranteeing, the possibility of modeling spatial and non-manual grammar.

This transcoding tendency is not merely architectural; it can be related to what Bauman (2008) identified as the phonocentric structure of Western linguistics. However, for the purposes of this analysis, the critical insight is not only philosophical but translational: text-centric representations entail identifiable translational loss, as computational work on glosses itself acknowledges (Müller et al., 2023). When a gloss-based system encounters a complex ASL construction—such as a classifier predicate that simultaneously encodes the shape, location, movement, and spatial relationship of multiple referents through a single manual configuration combined with specific facial grammar—it cannot fully represent this construction in the gloss. The gloss typically reduces it to a single lexical item, stripping away the simultaneous morphological information that constitutes the semantic core of the utterance. What is lost is not peripheral information but meaning central to the utterance.

The text-to-sign generation pipeline displays this transcoding logic in reverse. Many current generation systems decompose the process into stages such as text parsing, gloss generation, and skeletal pose animation, although some recent models map text to pose directly (Saunders et al., 2020). The signed output is largely a realization of a textual script, with limited capacity for the creative, context-sensitive, and culturally inflected choices that characterize human sign language interpreting. As interdisciplinary reviews of the field note (Bragg et al., 2019), the avatar's facial expressions are often absent or rudimentary, its use of signing space is geometrically simplified, and its register is typically flat. When such output is offered as a media accessibility service, automation risks scaling a counterfeit of access rather than access itself (Galasso, 2026b). From a TS perspective, this can be described, with Baker (1992), as a shortfall in pragmatic equivalence: the system may produce morphologically plausible output that is communicatively weak. Vermeer's (1989/2012) skopos theory further illuminates the problem: without a defined communicative purpose anchored in the needs of Deaf target users, these systems tend to optimize for engineering metrics (pose accuracy, joint-angle error) rather than for translational adequacy.

The hegemony of the BLEU metric (Papineni et al., 2002) illustrates this text-centric impasse most clearly. As documented in section 4.6, BLEU scores served as the primary evaluation metric in 72.4% of the reviewed studies. BLEU was designed to measure n-gram overlap between machine-generated text and a reference translation—a purely lexical, surface-level comparison that presupposes that translation quality can be reduced to string-matching. As a measure of lexical

overlap in written output, it has a legitimate role, and in sign-to-text evaluation it compares two written sentences; the difficulty arises when it serves as the main proxy for translation quality, since the written reference cannot register what the signed source conveyed through space and non-manual channels. Applying such a metric in this way to sign language machine translation is the point at which text-centrism is most visible in evaluation: it forces the assessment of a multimodal, visuo-spatial, embodied communicative act into the mold of a text-to-text comparison, rendering invisible precisely those dimensions—spatial grammar, non-manual markers, classifier predicates, use of signing space—that constitute the grammatical and pragmatic core of signed languages. The consequence is not merely methodological but epistemological: by optimizing for BLEU, SLMT research risks pursuing *textual mimicry* rather than translational quality, training models to produce outputs that satisfy a text-based metric without evidence that they meet the communicative needs of Deaf users. From a Translation Studies perspective informed by Vermeer's (1989/2012) skopos theory, this amounts to optimizing for a displaced *skopos*: the purpose being served is not the communicative adequacy of the translation for its intended audience but the numerical performance of the system on an evaluation benchmark that encodes text-centric assumptions. The near-total absence of human evaluation by Deaf participants (only 3.4% of studies) indicates that metric optimization has largely taken the place of user-centered quality assessment. What is urgently needed are evaluation frameworks grounded in TS principles—incorporating pragmatic adequacy ratings, communicative effectiveness testing, and systematic assessment by Deaf translators and users—that can capture the multimodal, cultural, and embodied dimensions that BLEU structurally excludes.

5.3 What Translation Studies Can Offer: Four Contributions

It is proposed that Translation Studies can contribute to the SLMT field in at least four concrete ways. First, TS's theorization of *equivalence* (Nida, 1964; Baker, 1992; Pym, 2010)—particularly the distinction between formal equivalence (surface-level correspondence) and dynamic equivalence (communicative effect on the target audience)—can reframe the evaluation crisis in SLMT. A TS-informed evaluation framework would privilege comprehension testing and communicative adequacy ratings by Deaf users over automated n-gram matching.

Second, *skopos theory* (Vermeer, 1989/2012) insists that the purpose of a translation determines its strategies and quality criteria. Different communicative contexts—medical consultations, legal proceedings, and educational settings—demand fundamentally different translation approaches. Current SLMT systems are largely context-blind, trained on decontextualized sentence pairs. A skopos-informed approach would argue for domain-specific SLMT systems calibrated to specific situations.

Third, the multimodal turn in TS, exemplified by research in audiovisual translation and accessibility studies (Asimakoulas, 2024; Díaz-Cintas &

Remael, 2021; Gambier, 2006), provides analytical tools for understanding how meaning is constructed across visual, spatial, temporal, and kinetic channels simultaneously—precisely what is missing from SLMT systems that reduce signing to a sequence of discrete glosses or skeletal keypoints.

Fourth, corpus-based TS methodologies, as implemented in COPA-TRAD (Silva & Fernandes, 2020), offer models for building parallel corpora that systematically compare human sign language interpretation with AI-generated translations. The creation of sign language parallel corpora annotated with TS-informed metadata—capturing not only lexical glosses but also discourse structure, register, pragmatic function, non-manual grammatical markers, and use of signing space—would constitute a transformative resource for both TS researchers and AI developers. Such corpora would enable the systematic identification of patterns of mistranslation, pragmatic loss, cultural flattening, and register misalignment in SLMT systems, analyses that are currently impossible because no such resources exist in the TS tradition. For Libras, a methodological route toward such resources has already been proposed, combining pose estimation, transformer-generated suggestions of lexical labels, and validation by Deaf linguists in a human-in-the-loop cycle, precisely to avoid the loss of articulatory information that glossing entails (Galasso, 2026a).

5.4 Toward Intersemiotic AI Translation Studies: A Research Agenda

Based on the gaps identified in this review (see Figure 4), a research agenda is proposed for what is here termed *Intersemiotic AI Translation Studies* (IAITS)—a transdisciplinary field at the intersection of Translation Studies, artificial intelligence, Deaf Studies, and Accessibility Studies. This agenda comprises five interconnected priorities.

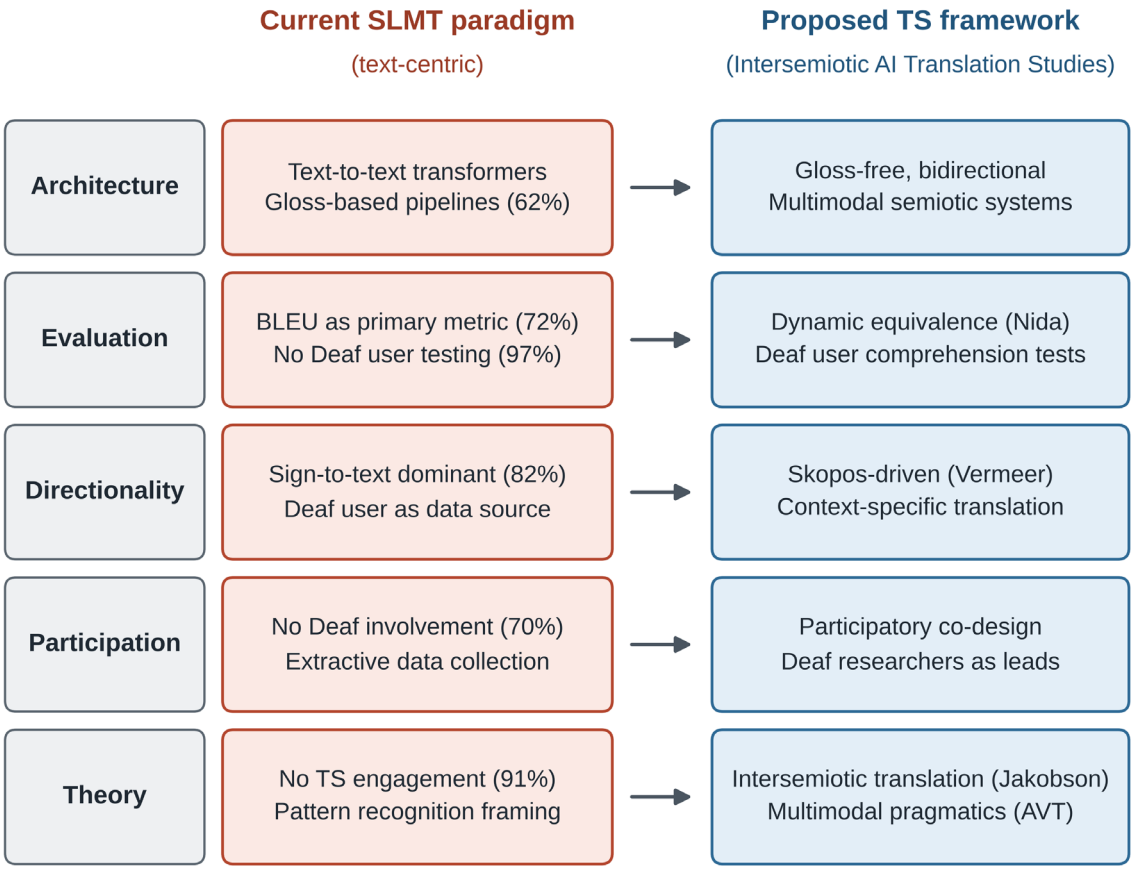


Figure 4. Epistemological gap map: Current SLMT paradigm vs. proposed Translation Studies framework.

First, the development of *multimodal evaluation frameworks* that assess SLMT quality beyond text-based metrics, incorporating pragmatic adequacy, cultural fidelity, and Deaf user comprehension. Second, the creation of *large-scale, linguistically annotated sign language parallel corpora* designed according to TS principles, with metadata capturing discourse context, communicative purpose (skopos), and pragmatic function. Third, the design of *gloss-free, bidirectional SLMT systems* that treat signed languages as complete semiotic systems rather than as visual encodings of spoken language. Fourth, the *mandatory inclusion of Deaf communities* as co-designers and co-evaluators, following participatory and emancipatory research methodologies (Kusters et al., 2017). Fifth, the *systematic cross-pollination* of TS and AI research to bridge the disciplinary gap documented in this review.

5.5 Implications for the Brazilian Context

This research agenda has particular relevance for Brazil, which offers uniquely favorable conditions for advancing IAITS. Libras (Brazilian Sign Language) was recognized as a legal means of communication and expression under Federal Law 10.436/2002, and Decree 5.626/2005 mandated its inclusion

in teacher education (Souza & Galasso, 2026). Law 14.191/2021 subsequently established bilingual Deaf education as an autonomous modality of schooling (Galasso, 2026c). Luso-Brazilian research teams have developed bilingual digital platforms for Deaf education, integrating computer vision, machine learning, and sign language video to create accessible educational content (Escudeiro et al., 2022). These platforms represent advanced experiments in applied intersemiotic AI translation, though they have not yet been theorized as such within Translation Studies. The intersection of Brazil's progressive legal framework, its vibrant TS tradition (exemplified by journals such as *Cadernos de Tradução* and *Ilha do Desterro*, and by research infrastructure such as COPA-TRAD), and its growing AI research capacity makes Brazil uniquely positioned to advance this transdisciplinary agenda.

5.6 Limitations

This review has several limitations. First, the search was limited to four databases and three languages, potentially excluding relevant studies in other languages or discipline-specific databases. Second, the categorization of studies according to TS concepts involves interpretive judgments. Third, the quantitative findings should be interpreted as indicative of trends rather than precise measurements. Fourth, the review maps published literature and may not capture ongoing or unpublished work, particularly from Deaf researchers who may face additional barriers to academic publication. Fifth, the charting was conducted by a single coder; although the operational definitions and decision rules are made explicit (Table 1), inter-rater reliability could not be estimated, and the categories that involve the most interpretation (translational framing, engagement with TS) are the most exposed to this limitation. Sixth, the interpretive reading proposed in section 5.1 is not tested by the charted data; it is a theoretically motivated account of them, which future work could examine empirically. Despite these limitations, the review provides a comprehensive and theoretically grounded map of the field that can serve as a foundation for the proposed IAITS agenda.

6. Conclusion

This scoping review has mapped 87 studies on AI-driven sign language machine translation published between 2019 and 2025, revealing a field characterized by rapid technological innovation but significant epistemological limitations. Empirically, the review documents a pronounced text-centrism, manifested in the predominance of gloss-based architectures, text-centric evaluation metrics, unidirectional translation priorities, the limited modeling of non-manual markers, and the marginal involvement of Deaf community members. Interpreted through Derrida (1967/1997) and Bauman (2008), this text-centrism represents not merely a technical challenge but what may be read as a structural bias consonant with the Western metaphysical tradition that Derrida

identified and that Bauman showed operates with particular force against signed languages.

Translation Studies, with its rich theoretical apparatus encompassing intersemiotic translation, equivalence theory, skopos theory, multimodal pragmatics, and corpus-based methodologies, offers indispensable conceptual tools for reframing SLMT as a problem of intercultural and intersemiotic mediation. At the same time, the challenge of SLMT pushes Translation Studies to confront its own limitations: its historical focus on written and spoken languages, its relative neglect of signed languages, and its insufficient engagement with computational methods and AI technologies.

The proposed agenda for Intersemiotic AI Translation Studies seeks to catalyze a genuine transdisciplinary dialogue between Translation Studies, computer science, Deaf Studies, and Accessibility Studies. Only through such dialogue can the field develop AI translation systems that truly serve the communicative needs and rights of Deaf communities—systems that respect signed languages not as impoverished approximations of 'real' language to be decoded into text, but as complete, sophisticated, and culturally rich semiotic systems deserving of translation on their own terms. The stakes of this endeavor are not merely academic: they concern the linguistic rights, social inclusion, and communicative autonomy of millions of Deaf people worldwide. As AI technologies continue to reshape the global translation landscape, Translation Studies has both an opportunity and a responsibility to ensure that this transformation does not reproduce the historical marginalization of signed languages. The concept of algorithmic logocentrism introduced in this article provides a critical lens through which researchers, developers, and policymakers can interrogate the hidden biases embedded in the very infrastructure of machine translation, working toward a more linguistically just and semiotically inclusive future.

Concretely, this article contends that the path toward genuinely translational AI systems for signed languages does not lie in deeper neural architectures or larger training corpora alone, but in a fundamental restructuring of the development pipeline around what may be termed a *translator-in-the-loop* paradigm. In current practice, SLMT systems are designed, trained, and evaluated predominantly by computer scientists and engineers, with Deaf communities most often positioned as data providers. The translator-in-the-loop model inverts this hierarchy: it positions Deaf translators, sign language linguists, and corpus-based Translation Studies researchers as integral participants at every stage of the AI development cycle—from the design of annotation schemas that capture non-manual markers and spatial grammar, through the construction of multimodal parallel corpora informed by TS principles, to the fine-tuning and evaluation of models against communicative adequacy criteria rather than text-based metrics. This is not merely a methodological recommendation but an epistemological imperative: as long as the knowledge systems of Deaf communities and the theoretical frameworks of Translation Studies remain external to the engineering process, SLMT systems are likely to keep reproducing the text-centric patterns

documented in this review, and the phonocentric assumptions they arguably instantiate. The transition from an extractive, engineering-driven paradigm to a participatory, translation-driven one represents the most consequential practical implication of the Intersemiotic AI Translation Studies agenda proposed here, and it is through this transition that the discipline of Translation Studies can most directly contribute to the development of AI technologies that serve linguistic justice.

DATA AVAILABILITY STATEMENT: Does not apply.

References

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. https://doi.org/10.1080/1364557032000119616

Asimakoulas, D. (2019). *Rewriting humour in comic books: Cultural transfer and translation of Aristophanic adaptations*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-19527-4

Asimakoulas, D. (2024). Audio described comics in the museum. *Meta*, 69(1), 174–197. https://doi.org/10.7202/1113946ar

Baker, M. (1992). *In other words: A coursebook on translation*. Routledge.

Baker, M. (1993). Corpus linguistics and Translation Studies: Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 233–250). John Benjamins.

Bauman, H.-D. L. (2008). Listening to phonocentrism with Deaf eyes: Derrida's mute philosophy of (sign) language. *Essays in Philosophy*, 9(1), 41–54. https://doi.org/10.5840/eip20089118

Berman, A. (2000). Translation and the trials of the foreign (L. Venuti, Trans.). In L. Venuti (Ed.), *The Translation Studies reader* (pp. 284–297). Routledge. (Original work published 1985).

Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., Caselli, N., Huenerfauth, M., Kacorri, H., Verhoef, T., Vogler, C., & Ringel Morris, M. (2019). Sign language recognition, generation, and translation: An interdisciplinary perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 16–31). ACM. https://doi.org/10.1145/3308561.3353774

Camgoz, N. C., Hadfield, S., Koller, O., Ney, H., & Bowden, R. (2018). Neural sign language translation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 7784–7793). IEEE. https://doi.org/10.1109/CVPR.2018.00812

Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10020–10030). IEEE. https://doi.org/10.1109/CVPR42600.2020.01004

Charlton, J. I. (1998). *Nothing about us without us: Disability oppression and empowerment*. University of California Press.

Derrida, J. (1997). *Of grammatology* (G. C. Spivak, Trans.). Johns Hopkins University Press. (Original work published 1967).

Desai, A., De Meulder, M., Hochgesang, J. A., Kocab, A., & Lu, A. X. (2024). Systemic biases in sign language AI research: A Deaf-led call to reevaluate research agendas. In *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources* (pp. 54–65). ELRA and ICCL.

Díaz-Cintas, J., & Remael, A. (2021). *Subtitling: Concepts and practices*. Routledge.

Dusi, N. (2015). Intersemiotic translation: Theories, problems, analysis. *Semiotica*, 2015(206), 181–205. https://doi.org/10.1515/sem-2015-0018

Escudeiro, P., Galasso, B., Teixeira, D., Gouveia, M. C., & Escudeiro, N. (2022). Inclusive MOOC – educational content for deaf people, a Portuguese proof of concept. *Cypriot Journal of Educational Sciences*, 17(4), 1269–1287. https://doi.org/10.18844/cjes.v17i4.7151

Galasso, B. (2026a). Corpus lexicográfico e visão computacional: Uma metodologia baseada em IA para a coleta e anotação semiautomática de sinais em larga escala e a análise de variações lexicais (regionalismos) em Libras [Lexicographic corpus and computer vision: An AI-based methodology for the large-scale semi-automatic collection and annotation of signs and the analysis of lexical variation (regionalisms) in Libras]. *Domínios de Lingu@gem*, 20, e020004. https://doi.org/10.14393/dlv20a2026-4

Galasso, B. (2026b). Generative AI and media accessibility for disabled audiences in Ibero-America. *Cuadernos.info*, (65), 168–188. https://doi.org/10.7764/cdi.65.106244

Galasso, B. (2026c). A LDB e a conquista da modalidade bilíngue: Trinta anos de tensões entre a integração, a inclusão e o horizonte da acessibilidade digital [The LDB and the achievement of the bilingual modality: Thirty years of tensions between integration, inclusion, and the horizon of digital accessibility]. *Dialogia*, (56), e30501. https://doi.org/10.5585/56.2026.30501

Galasso, B. (2026d). A prótese e o silêncio: Cognição estendida, ouvintismo algorítmico e injustiça epistêmica na era da inteligência artificial [The prosthesis and the silence: Extended cognition, algorithmic audism and epistemic injustice in the age of artificial intelligence]. *Revista de Filosofia Aurora*, 38, e202634314. https://doi.org/10.1590/2965-1557.038.e202634314

Galasso, B. (2026e). Visuo-spatial cognition and the redesign of human-AI interaction in generative artificial intelligence. *Discover Artificial Intelligence*, 6(1), 950. https://doi.org/10.1007/s44163-026-02040-2

Gambier, Y. (2006). Multimodality and audiovisual translation. In M. Carroll, H. Gerzymisch-Arbogast, & S. Nauert (Eds.), *Audiovisual translation scenarios: Proceedings of the Marie Curie Euroconferences MuTra* (pp. 1–8). EU-High-Level Scientific Conference Series. https://euroconferences.info/proceedings/2006_Proceedings/2006_Gambier_Yves.pdf

Gottlieb, H. (2005). Multidimensional translation: Semantics turned semiotics. In H. Gerzymisch-Arbogast & S. Nauert (Eds.), *Proceedings of the Marie Curie Euroconferences MuTra: Challenges of Multidimensional Translation* (pp. 1–29). EU-High-Level Scientific Conference Series. https://euroconferences.info/proceedings/2005_Proceedings/2005_Gottlieb_Henrik.pdf

Jakobson, R. (2000). On linguistic aspects of translation. In L. Venuti (Ed.), *The Translation Studies reader* (pp. 113–118). Routledge. (Original work published 1959).

Kusters, A., De Meulder, M., & O'Brien, D. (Eds.). (2017). *Innovations in Deaf Studies: The role of Deaf scholars*. Oxford University Press.

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science*, 5(1), 69. https://doi.org/10.1186/1748-5908-5-69

Lommel, A. R., Uszkoreit, H., & Burchardt, A. (2014). Multidimensional quality metrics (MQM): A framework for declaring and describing translation quality metrics. *Tradumàtica*, 12, 455–463. https://doi.org/10.5565/rev/tradumatica.77

Müller, M., Jiang, Z., Moryossef, A., Rios, A., & Ebling, S. (2023). Considerations for meaningful sign language machine translation based on glosses. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers) (pp. 682–693). ACL. https://doi.org/10.18653/v1/2023.acl-short.60

Nida, E. A. (1964). *Toward a science of translating: With special reference to principles and procedures involved in Bible translating*. E. J. Brill.

OpenAI. (2023). *GPT-4 technical report*. https://doi.org/10.48550/arXiv.2303.08774

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). BLEU: A method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics* (pp. 311–318). ACL. https://doi.org/10.3115/1073083.1073135

Pérez-González, L. (2014). *Audiovisual translation: Theories, methods and issues*. Routledge.

Pfau, R., Steinbach, M., & Woll, B. (Eds.). (2012). *Sign language: An international handbook*. De Gruyter Mouton.

Pym, A. (2010). *Exploring translation theories*. Routledge.

Quadros, R. M. de, & Segala, R. R. (2015). Tradução intermodal, intersemiótica e interlinguística de textos escritos em português para a Libras oral [Intermodal, intersemiotic and interlinguistic translation of written Portuguese texts into oral Libras]. *Cadernos de Tradução*, 35(2), 354–386. https://doi.org/10.5007/2175-7968.2015v35nesp2p354

Saunders, B., Camgoz, N. C., & Bowden, R. (2020). Progressive transformers for end-to-end sign language production. In *Computer Vision – ECCV 2020: Proceedings of the European Conference on Computer Vision* (pp. 687–705). Springer. https://doi.org/10.1007/978-3-030-58621-8_40

Silva, C. E., & Fernandes, L. (2020). Apresentando o COPA-TRAD versão 2.0: Um sistema com base em corpus paralelo para pesquisa, ensino e prática da tradução [Introducing COPA-TRAD version 2.0: A parallel corpus-based system for translation research, teaching and practice]. *Ilha do Desterro*, 73(1), 297–316. https://doi.org/10.5007/2175-8026.2020v73n1p297

Souza, I. S., & Galasso, B. (2026). A Libras como língua autônoma no Brasil: Uma análise discursiva dos marcos legais e das tensões na educação bilíngue de surdos (1988–2023) [Libras as an autonomous language in Brazil: A discourse analysis of

legal frameworks and tensions in bilingual Deaf education (1988–2023)]. *Revista DCS*, 23(91), e5892. https://doi.org/10.54899/dcs.v23i91.5892

Toury, G. (2012). *Descriptive Translation Studies and beyond* (revised ed.). John Benjamins.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. https://doi.org/10.7326/M18-0850

United Nations. (n.d.). *International Day of Sign Languages: 23 September*. Retrieved November 3, 2026, from https://www.un.org/en/observances/sign-languages-day

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems 30* (pp. 5998–6008). Curran Associates.

Venuti, L. (1995). *The translator's invisibility: A history of translation*. Routledge.

Vermeer, H. J. (2012). Skopos and commission in translational action. In L. Venuti (Ed.), *The Translation Studies reader* (3rd ed., pp. 191–202). Routledge. (Original work published 1989).

Yin, K., Moryossef, A., Hochgesang, J., Goldberg, Y., & Alikhani, M. (2021). Including signed languages in natural language processing. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (Volume 1: Long Papers) (pp. 7347–7360). ACL. https://doi.org/10.18653/v1/2021.acl-long.570

Submission date: 16/03/2026
Acceptance date: 09/10/2026
Editors: Lincoln P. Fernandes, Dimitris Asimakoulas, Ewerton Gleison Lopes Branco