technologies

MDPI

*Article*

# A Comparative Framework for Formal Representation Strategies in Sign Language Avatar Systems

**Nurzada Amangeldy [1,2,3], Aigerim Yerimbetova [1,4,*], Marek Milosz [1,5], Akmaral Kassymova [6], Elmira Daiyrbayeva [1,7] and Nazira Tursynova [1,2]**

[1] Institute of Information and Computational Technologies of the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan, Almaty 050000, Kazakhstan; amangeldi_n_3@enu.kz (N.A.); m.milosz@pollub.pl (M.M.); e.daiyrbayeva@satbayev.university (E.D.); ntursynova000@gmail.com (N.T.)
[2] Faculty of Information Technologies, L.N. Gumilyov Eurasian National University, Astana 010008, Kazakhstan
[3] SignBridge LLP, Astana 010000, Kazakhstan
[4] School of Engineering and Information Technology, META University, Almaty 050012, Kazakhstan
[5] Department of Computer Science, Lublin University of Technology, 20-618 Lublin, Poland
[6] Institute of Economics, Information Technologies and Professional Education, Zhangir Khan West Kazakhstan Agrarian-Technical University, Uralsk 090000, Kazakhstan; kasimova.akmaral2020@gmail.com
[7] Department of Software Engineering, Satbayev University, Almaty 050010, Kazakhstan
* Correspondence: aigerian8888@gmail.com; Tel.: +7-7015833302

**Abstract**

This paper proposes a unified methodological framework for evaluating heterogeneous approaches to avatar-based sign language visualization. The study introduces a four-dimensional analytical framework based on four independent criteria: (A1) pipeline architecture and degree of automation, (A2) data and annotation requirements, (A3) portability across sign languages and domains, and (A4) integration and accessibility. The framework is applied to a comparative analysis of three dominant paradigms: (P1) notation → animation (e.g., HamNoSys), (P2) writing-based representation → animation (e.g., SignWriting), and (P3) keypoint-based animation and Artificial Intelligence (AI) methods. The comparative assessment shows that the differences between the paradigms are structural and reflect trade-offs among linguistic accuracy, automation level, scalability, and user accessibility, rather than the superiority of any one technology. Overall, the structured comparative framework (A1–A4) is applied for analyzing three paradigms of sign language avatar generation. It enables a systematic evaluation of architectural, data-related, and practical characteristics, highlighting key trade-offs between linguistic accuracy, scalability, and accessibility.

**Keywords:** sign languages; HamNoSys; SiGML; SignWriting; SWML; avatar; glosses; multilingual translation; real-time; accessibility

Check for updates

Academic Editor: Juan Gabriel Avina-Cervantes

Received: 15 April 2026
Revised: 6 May 2026
Accepted: 11 May 2026
Published: 14 May 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

The World Health Organization estimates that approximately 430 million people worldwide have hearing loss, including more than 70 million Deaf individuals [1]. International organizations highlight the importance of sign language for ensuring equal access to information, communication, and fundamental rights for the Deaf community [2]. Sign languages are autonomous, natural languages with their own complex grammatical and syntactical structures, serving as the primary mode of communication and the foundation

of cultural identity for the Deaf community. Without the possibility of native sign language, Deaf people do not have equal rights and opportunities in society. Ensuring access to sign language is therefore a necessary condition for social participation and equality.

However, in the digital age, Deaf communities face serious disparities in access to information and technology. Most online content and digital services are targeted at hearing users and are presented in audio or written form, making them inaccessible to many Deaf people. Even the presence of subtitles and text transcriptions does not completely solve the problem, as a significant proportion of Deaf individuals have difficulty understanding written text—for many of them, it is a second language not directly related to sign language [3].

Lexicographic resources for sign languages are much smaller and organized differently from those for spoken languages. For example, ASL-LEX documents about 1000 ASL signs, while comparable projects on spoken languages cover tens of thousands of lexemes [4]. For Chinese Sign Language, the basic list of 902 gestures covers about 77% of usage, indicating a compact 'core' frequency [5]. Sign languages rely on a higher proportion of iconic units compared to English and Spanish [6]. Even with literacy in written language, the scarcity of sign language vocabulary limits understanding. Studies show a direct relationship between the volume of the dictionary and the understanding of the text in Deaf students (even those who can read) [7], and in Deaf students, the vocabulary and their 'knowledge of the world' are statistically lower than in hearing peers [8]. The risk of language deprivation in childhood leads to long-term cognitive and academic consequences, which underscores the importance of early and guaranteed access to sign language [9].

As a result, millions of Deaf people are disconnected from vital information and services. For example, during the COVID-19 pandemic, many Deaf people felt cut off from everyday life and had difficulty accessing remote medical counseling and reliable health information, as these materials were rarely provided in sign language [10]. This situation represents a form of "digital injustice" in which a large part of the population is deprived of equal access to digital resources and communication on the basis of language.

The reasons for this digital inequality are related to both sociolexical [4–10] and technical factors. The development of modern technologies traditionally took place without taking into account the needs of the Deaf. Many digital products and services are designed with an audist bias (focusing solely on hearing users) and without sufficient involvement of Deaf people themselves. The researchers note that this approach leads to the creation of tools that are not suitable for the real needs of end users [11].

One of the main reasons for technical barriers is the lack of data for developing recognition and translation algorithms for sign languages.

Despite the presence of separate initiatives, large, open, and linguistically aligned (video + gloss) datasets suitable for training modern SLR/SLT remain rare. The main resources are focused on ASL [11–14], several European sign languages [15–18], and a number of Asian sign languages [5,19,20]. For many other national sign languages, including those used by large communities, there are no datasets of comparable scale and openness, or the available datasets are limited in accessibility and annotation quality.

As a result, even advanced Artificial Intelligence (AI) systems for working with languages are far behind in the field of support for sign languages. All of this is exacerbating the digital accessibility gap, in which Deaf communities around the world are left out of the benefits of the information society. It is necessary to recognize this issue as a matter of social justice and to draw the attention of researchers and developers to the task of ensuring the linguistic accessibility of digital technologies for Deaf users at the global level.

An important approach to addressing digital accessibility is the development of sign language avatar systems. These systems transform written or spoken input into animated

signs via multi-stage pipeline topologies. An essential component of these pipelines is the utilization of intermediate representation layers, including formal sign notations, writing-based representations, or skeletal keypoints that abstract linguistic content prior to motion generation. Examining these pipeline structures, especially their intermediary representation layers, can yield significant insights into the trade-offs among automation, language purity, and accessibility.

Recent evaluations have offered significant classifications of sign language avatar technologies, recognition techniques, and representation frameworks. The surveys have delineated the domain of sign synthesis methodologies, encompassing rule-based and neural animation [21,22], gloss-to-motion frameworks [23,24], and multimodal input techniques that integrate visual and sensor inputs [25–28]. Certain studies have examined formal notation systems such as HamNoSys and SignWriting (SW) [29], while others investigate accessibility and educational applications for Deaf individuals [30]. These studies have substantially illuminated advancements, constraints, and prospective trajectories in avatar design and simulation production pipelines.

Nevertheless, the majority of these studies are predominantly descriptive and lack a cohesive architectural viewpoint. They document current systems and technologies, but infrequently define the pipeline architectures and intermediate representation layers involved in sign creation processes. The absence of a systematic framework for comparing the encoding, transformation, and rendering of sign content across various techniques constrains a comprehensive study. This study does not aim to provide a comprehensive survey; instead, it introduces a structured methodology for comparing the intermediate representation layers underlying the three principal paradigms of avatar-based sign language visualization. This facilitates a systematic assessment of trade-offs in automation, scalability, and accessibility in both present and prospective systems.

To overcome this gap, this study compares three paradigms of sign language visualization. A four-dimensional framework (A1–A4) is introduced to systematically analyze these approaches, applied to three paradigms: (1) notation → animation (HamNoSys → SiGML → JASigning), (2) writing-based representation → animation (SW → SignWriting Markup Language (SWML) → tuniSigner), and (3) keypoint approach (extraction and generation of poses through machine learning). These paradigms fundamentally differ in their pipeline topologies, each employing a distinct intermediate representation of sign content. For clarity, the four evaluation criteria are designated as A1–A4.

The comparison is based on four independent criteria: (1) the architecture and degree of automation, (2) data and annotation, including the presence of video + gloss pairs, (3) portability across languages and domains, and (4) integration and accessibility, covering API, web embedding, and offline. A1 pertains to pipeline architecture and automation; A2 addresses data and annotation needs; A3 concerns portability across languages and domains; and A4 relates to integration and accessibility.

**RQ1.** *What are the fundamental differences between the main paradigms of implementing sign avatars in terms of pipeline architecture and automation level?*

**RQ2.** *How do data requirements and annotations affect the scalability and applicability of different sign language visualization techniques?*

**RQ3.** *To what extent do different methods ensure portability between sign languages and subject areas?*

**RQ4.** *What trade-offs between linguistic accuracy, accessibility, and technical complexity are revealed when comparing existing avatar technologies?*

## 2. Materials and Methods

This paper presents a unified four-dimensional analytical framework for the comparative analysis of approaches to automatic sign language visualization based on formal representations and motion data. The main goal of the framework is to identify conceptual differences between approaches that determine the architecture, scalability, and practical application of sign language avatar systems. It is important to emphasize that these paradigms are considered not as competing implementations, but as different methodological strategies, each of which occupies its own position in the space of design solutions, defined by the criteria (A1–A4).

This study employs a formal comparative analysis that analyses current systems in terms of their architectural structure. The comparison utilizes both academic literature and publicly accessible documentation of current systems, including JASigning, tuniSigner, and MediaPipe, facilitating repeatability and independent validation. The comparison follows a structured evaluation protocol based on the previously defined criteria (A1–A4).

As a result of the analysis of the scientific literature [22,31–33] and existing solutions [34–38], it was found that most of the methods in the field of sign language visualization can be combined into three main paradigms. These paradigms are distinguished by their representation and visualization strategies (see Figure 1).

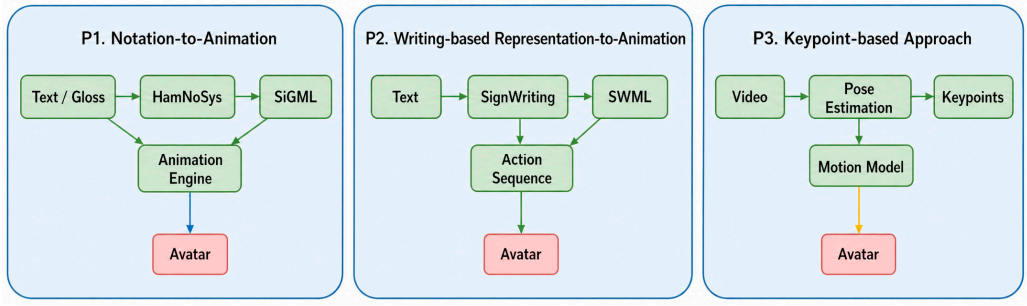



**Figure 1.** Paradigms of sign language visualization.

The choice of these three paradigms is predicated on their unique representation methodologies, including symbolic notation, writing-based representations, and data-driven motion representations. Each corresponds to a broader category of existing sign language technologies. This framework provides a balanced representation of rule-based, hybrid, and data-driven approaches. Each paradigm is further illustrated through representative case-study systems described in the literature, enabling a structured comparison of their architectural and operational characteristics (see Figure 2).

Each criterion represents a fundamental dimension that influences system performance and deployment. A1 evaluates the structural complexity and automation of the pipeline; A2 pertains to the scale and nature of necessary training data or symbolic annotation; A3 examines the system's adaptability to new languages and domains; and A4 focuses on integration into practical applications such as APIs, offline access, and web embedding, among others.

The set of selected criteria provides a comprehensive assessment of sign language visualization paradigms, covering linguistic, data-related, architectural, and operational aspects.

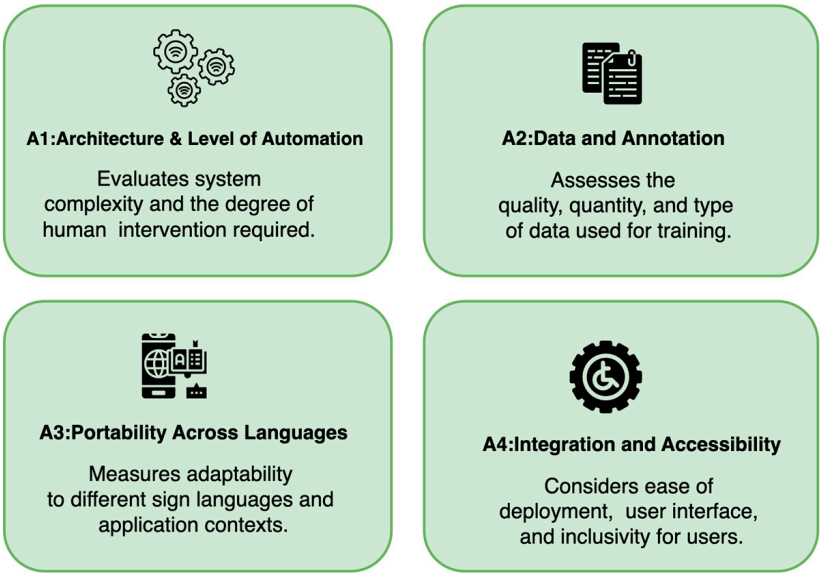



**Figure 2.** Four-dimensional analytical framework (A1–A4).

### *2.1. Analytical Framework*

#### 2.1.1. P1 Notation-to-Animation Systems

A1: In the P1 paradigm, the gesture is specified as a formal HamNoSys recording (notation), which is automatically translated into the SiGML gesture markup language [39,40], after which a special engine, such as JASigning, plays the gesture on a virtual avatar [41,42]. This pipeline enables automated motion generation from symbolic descriptions.

In P1 systems, the gesture is represented as a formalized configuration, including a set of interrelated manual and non-manual parameters. For two-handed gestures, a symmetry operator is used to copy or mirror the description of the dominant hand to the non-dominant hand. The initial state of the gesture is set through the hand configuration. This configuration includes a hand shape describing the mutual position of the fingers, the orientation of the hand, and a localization indicating the position of the hand relative to the body or gesture space. The dynamic component of the gesture is described by means of motion parameters that determine the trajectory and nature of the movements of the hands during the gesture execution. The notation description includes non-manual elements, such as head movements, direction of look, facial expressions, and articulation of the mouth. Together, these components form a complete formal representation of the gesture, which is used by notation-oriented systems for the subsequent generation of the animation of the sign avatar (see Figure 3).

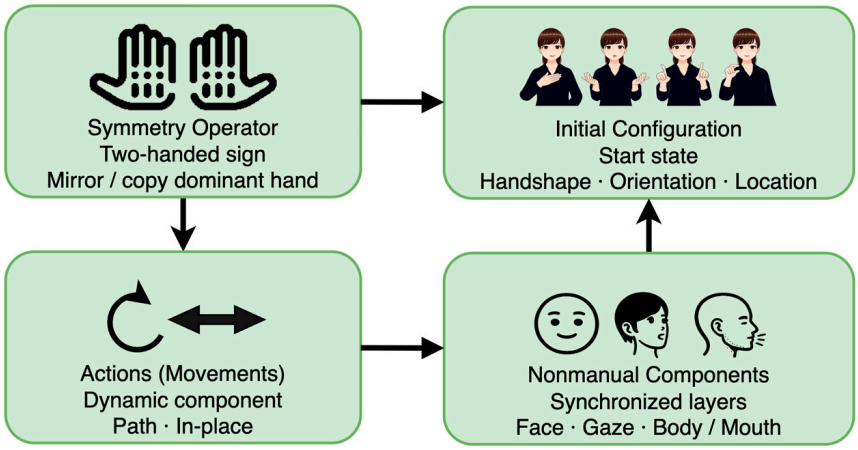



**Figure 3.** Structural components of HamNoSys-based sign representation.

Tools have been developed to implement such a pipeline, for example, the ESIGN editor contains a character that is based on ASL and BSL with ready-made HamNoSys/SiGML codes [43,44]. For new languages, it is enough to supplement the system with new HamNoSys records and convert them to SiGML, as shown in the example of the Indian Sign Language (ISL), for which a dedicated HamNoSys-to-SiGML conversion system with avatar output has been developed [41,43].

A2: The P1 approach requires a pre-vocabulary, where each gesture of the language must be described manually in the HamNoSys notation [45,46]. The notation itself is very detailed and accurate, but cumbersome to read and compile [43]. For example, HamNoSys allows the recording of facial expressions; the system of symbols for non-manual components is not yet fully completed and continues to expand. Creating annotated data in HamNoSys is time-consuming, but it gives a compact description of the movements instead of a large amount of video frames [44]. Given the absence of a universally accepted writing system for sign languages, symbolic representation becomes essential. Video recording of each gesture would require capturing multiple instances and contexts [47]. However, the collection and annotation of new data in HamNoSys is limited by the availability of experts and tools for this specific notation [48].

A3: A key advantage of P1 is the linguistic independence of HamNoSys [43,49]. This notation was originally developed as a universal system and is able to describe gestures from different sign languages, and does not rely on specific conventions. This means that the same animation system can be used for different national sign languages. It is enough to provide for each of them a vocabulary of gestures in HamNoSys. In practice, there are implementations for several languages, the most fully developed of which are the sign languages of India [41,43], the Arab region [42], Algerian [50] and the UK [51]. Portability between thematic domains is relatively high, since new vocabulary can be added without modifying the system architecture. To represent signs from new subject domains, the corresponding entries can be added to the system in HamNoSys or rendered through fingerspelling. Thanks to the expressive possibilities of notation, which describe the shape of the hands, orientation, localization, and type of movement, almost any new vocabulary can be covered without changing the architecture of the system.

A4: The P1 paradigm is relatively easy to integrate because it does not require heavy computation. The animation of signs is reduced to the reproduction of pre-encoded movements. Such a system can work in real time on conventional equipment, making it suitable for resource-constrained environments. Ready-made modules (e.g., JASigning library) have already been developed that can be integrated into software complexes for translating speech or text into gestures [41]. The avatars employed in this paradigm (refer to Figure 4) do not necessitate high visual fidelity. A fundamental 3D character, including articulated hands and a facial rig, adequately represents essential parameters, including handshape, orientation, location, and nonmanual markers. Systems such as JASigning currently provide SiGML-based animation and may be seamlessly integrated into educational or assistive apps with minimal hardware specifications.

The P1 paradigm, based on the HamNoSys → SiGML → Animation, demonstrates a high degree of formalization, modularity, and portability. Its architecture fully automates the generation of gesture animation with the lexicon in HamNoSys notation, and existing engines like JASigning provide the reproduction of these descriptions on virtual avatars [42].

The main limitation of the paradigm remains the dependence on manual annotation; that is, the creation and expansion of dictionaries requires the participation of experts familiar with the notation and structure of the sign language. However, it is this manual processing that provides linguistic accuracy and makes the system manageable. Due to the

language-independent nature of HamNoSys, this model is easily adapted to different sign languages and can be scaled to new domains by adding new units without rebuilding the entire architecture.



**Figure 4.** Examples of 3D signing avatars generated using HamNoSys-based notation-to-animation systems.

Thus, the P1 paradigm provides a flexible and reproducible framework for building multilingual sign systems. Despite the need for manual annotation, it remains a practical solution in scenarios where linguistic precision and deterministic generation are required for educational, information, and support systems of translation into sign language.

#### 2.1.2. P2 Writing-to-Animation Systems

A1: In the P2 paradigm, the input data is the recording of signing using the written SW system (specialized writing system developed for the Deaf community) [52,53]. This writing system captures gestures with graphical symbols placed in a two-dimensional "sign window" (see Figure 5). Special software, for example, the prototype tuniSigner, automatically parses such a recorded sign in SWML format (XML notation for SW) [54,55], extracts linguistic parameters (configuration and position of hands, movement, facial expressions), and based on these parameters forms a sequence of actions for the 3D avatar [18]. In other words, the static transcription of the gesture in the form of SW is transformed into a dynamic animation, where the virtual character performs recorded gestures without human intervention [38,53,56]. This is achieved through a multi-stage parsing of the record, the system classifies the SW symbols, determines their spatial location and sequence according to defined rules, and then generates the hand movement and facial expressions embedded in the original record [57].

A2: Unlike HamNoSys, SW was originally created for everyday use by the Deaf community and is already used to record many national sign languages [58,59]. This means that for the P2 paradigm, some existing resources are available, although the creation and expansion of datasets still rely heavily on manual input. Text corpora, dictionaries, and even literary works written by SW can be used in the system without laborious transcoding [46]. Moreover, SW has a rich set of graphemes (tens of thousands of characters), allowing subtle representation of gesture details—up to the expressive nuances of gesture poetry. Thanks to this flexibility, gesture annotation in SW provides a more complete and natural representation than highly specialized notations. However, complex algorithms are required to automatically interpret such records [34,52,53,60]. With more than 37,000 symbols distributed across hand configurations, movements, facial expressions, body parts, and spatial markers, ISWA provides a near-complete formal inventory for representing sign language gestures. Such a large and structured symbol set enables encoding of fine-grained phonological and expressive details, including contact, orientation, and dynamic features.

This richness potentially supports scalability and portability of writing-to-animation systems, as new signs and domains can be introduced solely by adding new SW records without modifying the underlying animation architecture.

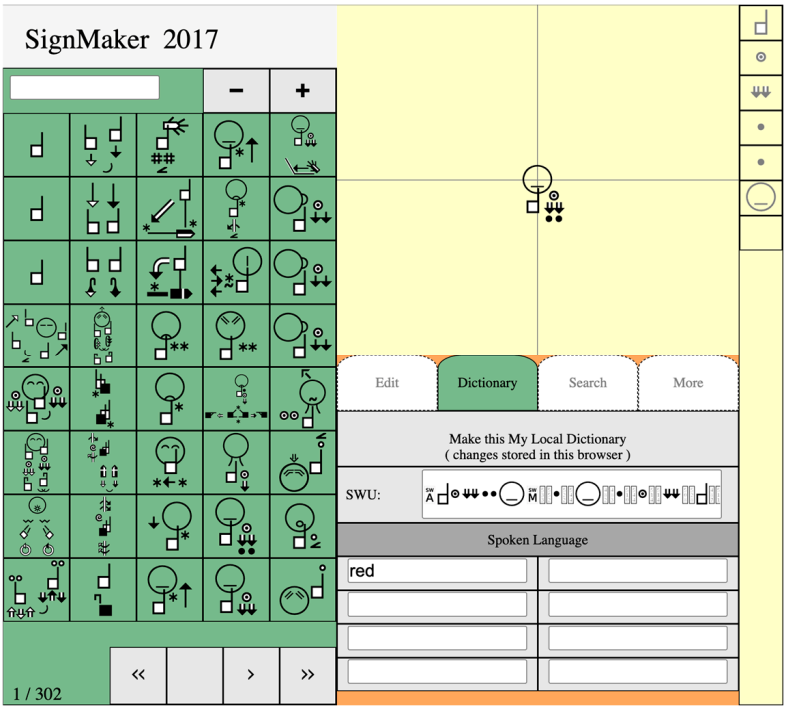



**Figure 5.** SignMaker 2017 interface for creating and editing SignWriting records in a two-dimensional sign window.

The SWML format captures only the set and position of the characters, without specifying a single order of actions, and the relationships between them can implicitly affect the meaning (for example, two-hand contact). Therefore, the P2 system needs to apply rules for interpreting the sign, for example, to determine which hand moves first, or to establish the fact of contact between the hands on the relative position of the symbols, when converting writing into animation. However, if the original records comply with the SW standard, they can be converted into gestures without additional manual markup. The advantage lies in the fact that the Deaf community itself can contribute to expanding such datasets. For Deaf users, learning SW is generally easier than mastering a specialist notation such as HamNoSys, which increases the scalability of data collection.

A3: SW was developed as a universal system and has already been adapted for many sign languages of the world, including Brazilian (Libras) [53,55], Swiss German [61,62], Indian [58,59], Myanmar, and Spanish. The ISWA alphabet that underpins it includes symbols for most possible configurations of hands, movements, and expressions, enabling representation of a wide range of gestures. This theoretically enables cross-language portability: one animation mechanism (for example, the tuniSigner engine) can interpret records in different languages, differing only in input data [63,64]. In principle, the architecture does not need to be modified when switching to a new language; it is sufficient to have an appropriate corpus or dictionary in SW for that language. Similarly, the method is adaptable to different subject domains. New concepts and terminology are introduced by adding their written equivalents (SW records) to the database [57,65]. Because the SW grapheme set is extensive, even specialized or technical gestures can usually be represented using existing symbols. Thus, the P2 paradigm theoretically offers portability similar to P1, although in practice its adoption remains limited, combining it with the benefits of a

true writing system, including the involvement of the Deaf community in the creation and expansion of resources.

A4: The main goal is to make sign language writing more accessible and practical for a wide audience of Deaf people [52,66]. Mastering SW requires special training, so the automatic visualization of such records significantly increases their practical usefulness [58,59,67–69]. A digital avatar that reproduces the written text allows the Deaf to receive information in their native language in sign form, without the need to interpret graphical symbols manually [52,66]. For example, children's stories or learning materials recorded in sign language in SW can be directly rendered by an animated avatar, making the learning and perception process more accessible [34,52,53,60]. Given that SW incorporates complex two-dimensional visual structures (see Figure 6), the avatar must accommodate a broad spectrum of movements and facial expressions. Despite the comparatively low computational demand, effective animation necessitates the avatar's ability to manage intricate spatial distinctions and symmetrical gesture patterns suggested by the text.

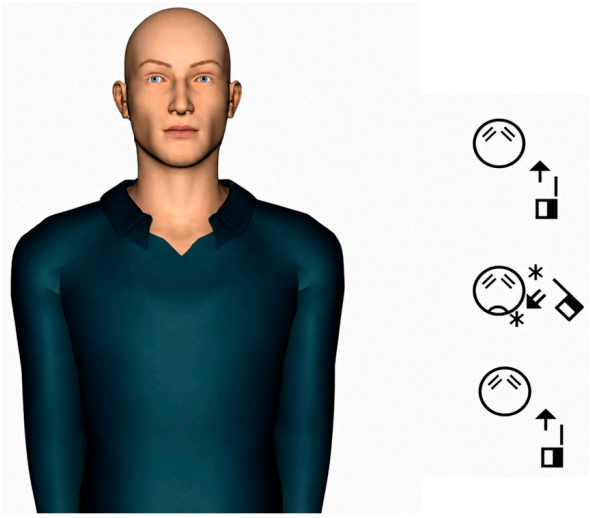


**Figure 6.** Example of a SignWriting-based dictionary entry with parallel written and avatar-based sign representation.

In technological terms, modules that interpret SWML and manage a 3D avatar have been developed for many applications, from eBooks to websites with sign language content [54,70,71]. Since gesture generation in P2 is based on rule conversion, the hardware requirements remain moderate, and browser-based deployment is possible with optimization. Thus, P2 solutions such as tuniSigner make the written form of sign language more accessible. They allow Deaf users to use their own writing system while enabling the automated transformation of visual-symbolic representations into dynamic sign language animations [38,56].

Thus, despite its high expressiveness and user-oriented nature, the Writing-to-Animation paradigm is limited by the low prevalence of written systems, the complexity of automatic analysis of two-dimensional recording, and the difficulties of integration, which reduces its applicability in scalable and automated sign avatar systems.

#### 2.1.3. P3 Keypoint-Based Animation Systems

A1: From an architectural point of view (A1), most systems following this paradigm use a multi-stage pipeline consisting of pose estimation, temporal motion simulation (see Figure 7), and avatar animation. Key points are usually obtained using computer vision-based pose assessment methods and are compared with skeletal keypoints and facial parameters to control the animation (see Figure 8). Studies on two-way sign language

communication systems show that pose estimation can serve as a reliable intermediate representation between sign language and avatar visualization, providing both recognition and synthesis [72]. Similarly, end-to-end systems that integrate recognition, translation, and gesture generation use key points as hidden motion descriptors before final rendering [73]. However, architectural solutions remain heterogeneous across studies, although the pipeline itself is typically fully automated, as studies do not use a standardized representation of motion.

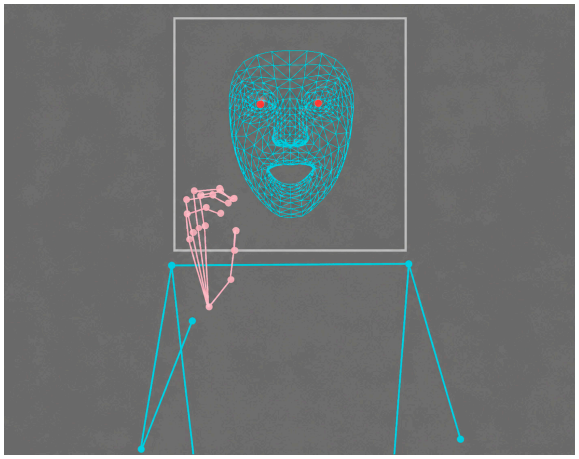


**Figure 7.** Key point–based motion representation used as an intermediate layer between pose estimation and avatar animation.

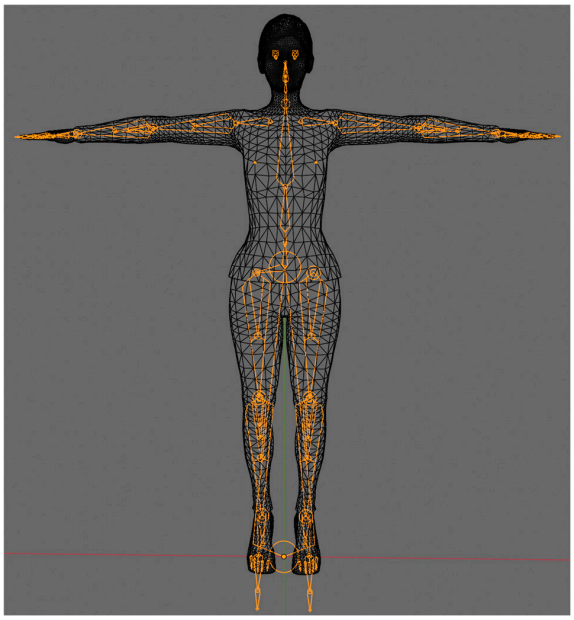


**Figure 8.** Keypoint-based skeletal representation used for language-independent motion encoding in sign language animation.

A2: With regard to data and annotation requirements (A2), key point-based approaches significantly reduce dependence on linguistic annotation. Most systems work with raw or slightly processed video data, extracting body and hand keypoints without the need for gloss-level or other manual linguistic annotation. This feature allows the reuse of large, unannotated video datasets and reduces the cost of creating datasets. Early work based on motion capture data treated skeletal joint trajectories as idealized key points, emphasizing their suitability for sign language animation, but also identifying scalability limitations due to specialized equipment requirements [74]. More recent computer vision approaches

eliminate some of these limitations, although they remain sensitive to detection errors, particularly in fine-grained finger articulation, as shown in comparative evaluations of key point identification techniques for avatar reconstruction [75].

A3: With regard to language-to-region portability (A3), the keypoint-based animation paradigm is structurally independent of language, since keypoints encode physical movement rather than linguistic units. This makes the approach potentially transferable to different sign languages and related areas, such as virtual reality, telepresence, and human–computer interaction [76]. However, a number of studies note that linguistic tolerance in practice remains limited. Differences in spatial grammar, language-specific non-manual components, and culturally conditioned sign conventions are not explicitly modeled when using only raw keypoint representations (see Figure 8) [72,73]. As a result, although movement transfer is technically feasible, linguistic adequacy may vary depending on the training data.

A4: Integration and accessibility (A4) are one of the advantages of this paradigm. Key point-driven animation systems can be integrated into real-time applications, including web platforms, virtual reality environments, and interactive avatars. Research in facial animation shows that mapping facial landmarks with blendshapes in real time is feasible and significantly improves perceived naturalness, which is crucial for sign language because of the grammatical role of non-verbal markers [77,78].

However, full integration of key points of the body, hands, and face significantly increases computational requirements and requires additional smoothing and synchronization mechanisms to prevent abrupt or unnatural transitions in motion. Within the proposed framework, this impacts A1 (pipeline complexity) and A4 (integration and accessibility). Accurate reproduction of fine-grained motion, including finger articulation and facial dynamics, is necessary but computationally demanding. As keypoint data is typically derived from real-world signing, the avatar (see Figure 9) must preserve motion consistency to support correct interpretation.



**Figure 9.** Example of a neutral 3D signing avatar used for gesture visualization and linguistic evaluation.

Overall, the keypoint-based paradigm offers a highly automated and scalable approach to creating sign language avatars, minimizing annotation requirements and providing real-time visualization. At the same time, the lack of standardized motion representations and explicit linguistic abstraction limits its interpretability and interlingual stability. These characteristics position this paradigm as an effective solution for realistic synthesis of movements, while emphasizing the need for additional mechanisms when linguistic precision and semantic control are required. Significantly, from a pragmatic standpoint, the keypoint-based paradigm provides an extra benefit in the end-user setting. Because most users are unfamiliar with formal notation systems such as HamNoSys or SW, keypoint-based systems provide a more accessible and transparent method for visualizing sign language.

Visual observation and machine extraction of gestures from real-world recordings enhance this paradigm's alignment with users' natural representations and require less specialized knowledge. This enhances the application of the technique in educational, communicative, and practical contexts, particularly in multilingual or cross-cultural settings.

### *2.2. Case Study Systems Overview*

To demonstrate the relevance of the proposed analytical approach, representative sign language avatar systems were selected from the literature. The selection was based on peer-reviewed publications and publicly documented implementations, guided by the following criteria: (1) coverage of the three paradigms (P1–P3), (2) availability of technical descriptions enabling analysis, and (3) relevance to recent developments in sign language visualization. These systems serve as case studies illustrating different technological approaches and enabling a structured comparison across paradigms.

The chosen systems vary in their representation methods, input modalities, and animation production approaches, providing a representative sample for comparative analysis based on the specified criteria (A1–A4) (refer to Table 1).

**Table 1.** Representative systems for the three sign language avatar paradigms (P1–P3).

| System | Paradigm | Input | Output | Source |
|---|---|---|---|---|
| Speech-to-Sign Language Interpreter | P1 | Speech | Avatar animation | [42] |
| Automatic Translation of Complex English Sentences to ISL | P1 | Text | Synthetic sign video | [45] |
| SW-based Machine Translation | P2 | Spoken language | SW/sign sequence | [61] |
| SW + Deep Learning Architecture | P2 | SW notation | Avatar animation | [35] |
| End-to-End Deep Learning Framework | P3 | Video | Generated sign video | [73] |
| Two-Way AI Communication System | P3 | Speech/sign video | Avatar communication | [72] |

The selected systems represent different technological approaches and reflect the core characteristics of each paradigm. Notation-based systems rely on formal linguistic representations such as HamNoSys or SiGML, writing-based representation approaches use visual-symbolic representation systems such as SW, and keypoint-based systems rely on computer vision and machine learning to produce motion directly from visual data.

These case studies provide a basis for assessing how each paradigm satisfies the analytical criteria (A1–A4) in terms of architecture, data requirements, portability, and integration.

## 3. Results

The application of the proposed analytical framework to the three dominant paradigms of sign language avatar generation highlights structural differences in automation, data requirements, portability, and system integration. The semi-quantitative comparative assessment based on the four analytical criteria (A1–A4) is presented in Table 2, while Table 3 provides a qualitative interpretation of these results, highlighting the key advantages and limitations associated with each paradigm.

**Table 2.** Quantitative evaluation of sign language avatar paradigms according to the four analytical criteria (A1–A4).

| Criterion | Scoring Interpretation (0–2) | P1 | P2 | P3 |
|---|---|---|---|---|
| A1: Architecture | 0—manual, no end-to-end pipeline; 1—partial automation with manual or complex steps; 2—fully or predominantly automated pipeline | 2 | 1 | 2 |
| A2: Data Requirements | 0—high reliance on expert/manual annotation; 1—mixed manual and automated data; 2—minimal annotation, data-driven learning | 0 | 0 | 2 |
| A3: Portability | 0—low, requires full redesign; 1—moderate with adaptation; 2—high with minimal changes | 1 | 0 | 2 |
| A4: Integration | 0—complex, requires specialized knowledge; 1—moderate integration effort; 2—easy integration, accessible to non-experts | 1 | 2 | 2 |

**Table 3.** Comparative analysis of sign language animation paradigms across four evaluation criteria (A1–A4).

| Paradigm | A1: Architecture | A2: Data Requirements | A3: Portability | A4: Integration | Advantages | Limitations |
|---|---|---|---|---|---|---|
| P1—Notation-to-animation | Rule-based; HamNoSys/ SiGML; fully automated. | Expert annotation; slow but structured. | Moderate; language-independent notation, but new lexicons needed. | Lightweight; real-time; limited for daily users. | Precise; modular; efficient on basic hardware. | Requires HamNoSys expertise; manual vocabulary expansion; rigid motion. |
| P2—Writing-to-animation | Rule-based parsing of SignWriting into SWML. | Community-driven; accessible; complex parsing. | Low; uneven global adoption of SignWriting. | Useful for education; difficult full software integration. | Intuitive; Deaf-user participation; community data creation. | Requires learning a special script; limited adoption; difficult 2D-to-3D conversion. |
| P3—Keypoint-based/AI | Data-driven; pose estimation; generative models. | Large video corpora; fewer linguistic annotations; error-sensitive. | High; keypoint representation is mostly language-independent. | Strong potential for web, VR, and telepresence. | Automated; scalable; natural motion; non-expert friendly. | Limited linguistic structure; high computing needs; corpus bias. |

The scoring scheme (0–2) reflects relative differences in structural properties across paradigms and is intended as an interpretative rather than statistical measure. The evaluation scores are grounded in the architectural structure of the three paradigms identified in the selected case study systems. Each paradigm follows a distinct motion generation pipeline that influences its level of automation, data requirements, portability, and integration capabilities.

The notation-to-animation paradigm (P1) relies on the formal linguistic encoding of signs using phonetic notation systems such as HamNoSys. The typical processing pipeline observed in the analyzed case studies can be summarized as follows: *Input Text* → *HamNoSys* → *SiGML Parser* → *3D Engine* → *Avatar*, where the symbolic description encoded in HamNoSys is converted into SiGML and interpreted by animation engines such as JASigning. This deterministic conversion process enables a high level of pipeline automation, which explains the assigned score for the architecture criterion (A1 = 2).

However, constructing HamNoSys lexicons requires extensive manual work by trained linguists, who must encode the phonological structure of each sign. This strong dependence on expert annotation leads to a low score for data and annotation requirements (A2 = 0).

Although HamNoSys is linguistically neutral and theoretically applicable across multiple sign languages, each language requires a dedicated lexical database, resulting in moderate portability (A3 = 1). In terms of integration, notation-to-animation systems are computationally efficient but require developers to understand HamNoSys notation in order to extend the vocabulary, which explains the moderate integration and accessibility score (A4 = 1).

The writing-to-animation paradigm (P2) uses visual symbolic representations of sign languages, most prominently SW. The processing pipeline identified in the analyzed systems can be described as follows: *Input Text* → *SW* → *SWML Parser* → *Action Sequence* → *Avatar.*

In this architecture, signs are encoded using two-dimensional graphical symbols representing handshape, orientation, and spatial relationships. These symbols are typically stored in markup formats such as SWML. To generate avatar animation, the symbolic representation must be parsed and translated into an ordered sequence of actions that the animation engine can interpret.

The transformation from a visual-symbolic two-dimensional representation into three-dimensional animated motion introduces additional algorithmic complexity. For this reason, writing-to-animation systems receive an intermediate score for pipeline automation (A1 = 1).

The creation of SW datasets also relies on manual input from annotators or community members, which limits scalability and results in a low score for data and annotation requirements (A2 = 0). Furthermore, the global adoption of SW remains uneven across regions, constraining cross-language portability and resulting in a low portability score (A3 = 0).

At the same time, the visual nature of SW makes it relatively intuitive for Deaf users and accessible through graphical interfaces, which contributes to a high integration and accessibility score (A4 = 2).

The keypoint-based paradigm (P3) relies on computer vision and machine learning to extract motion representations directly from visual input. The general pipeline observed in the analyzed systems can be summarized as follows: *Input Video* → *Pose Estimation* → *Keypoints* → *Motion Model* → *Avatar*.

In this paradigm, pose estimation algorithms detect skeletal keypoints of the body, hands, and face from video input. These landmarks are then processed by motion models, which may include deep neural network models, sequence models, or generative architectures that synthesize avatar motion.

Because the pipeline operates directly on motion data and does not require symbolic linguistic encoding, it achieves a high degree of automation, which explains the high A1 score (2). In addition, keypoint-based approaches can learn directly from raw video data, which significantly reduces the need for linguistic annotation. This capability justifies the highest score for data and annotation requirements (A2 = 2).

Since keypoint representations encode the physical movement of signing rather than language-specific symbolic structures, the approach can be adapted to different sign languages with relatively limited structural modification. Consequently, it receives the highest portability score (A3 = 2).

Finally, because users do not need to learn specialized notation systems, keypoint-based systems are more transparent for non-expert users and easier to integrate into modern AI pipelines, which explains the high integration and accessibility score (A4 = 2).

Overall, the results show that there are clear trade-offs between the three paradigms in terms of linguistic accuracy, user accessibility, scalability, and technical complexity. These differences highlight the importance of carefully choosing a sign language visualization approach based on the specific application, available resources, and target audience.

## 4. Discussion

This work focuses on the fact that choosing a specific technology to create avatars is not a search for the "best" solution, but a search for a balance between available resources and the required quality. The structural differences between the three paradigms (P1, P2, and P3) reflect systemic compromises and highlight the following key points:

Analysis of the literature shows that there was previously a significant gap in the systematic comparison of technologies. Attar et al. note that so far, no qualitative comparative analysis of gesture generation methods has been conducted [46]. Similarly, Aziz and Othman emphasize the lack of a standardized framework for avatar development [21]. This study addresses this gap by offering a unified framework (A1–A4) that allows for structured comparison of heterogeneous approaches (from HamNoSys to AI techniques) based on reproducible criteria rather than just bibliometric metrics.

Many modern AI-based (P3) systems achieve strong performance only when large-scale video datasets are available, such as American Sign Language [12–14,79]. For most national sign languages, such data are not available, which creates a barrier to accessing information [80]. In such circumstances, the use of rule-based systems (P1) [31–34] or writing-based representation approaches (P2) [53,69] becomes a more practical solution, as they allow content generation without requiring large-scale video data collection. This study highlights the SW (P2) paradigm as the most user-centered, allowing Deaf users to participate in data creation themselves. This echoes the findings of Aziz and Othman about the need for community involvement in the assessment process and the comments of the authors of the 2025 review that usability studies with Deaf users remain limited [21]. Although translating 2D symbols into 3D animation is technically difficult, this method is closest to the culture and needs of the Deaf community, making technology more user-centered [34,52,53,60]. The proposed framework takes into account the accessibility criterion (A4), which makes it a practical "roadmap" for developers when choosing an architecture depending on social and technical goals.

In this context, keypoint-based systems (P3) provide a distinct advantage: they do not require knowledge of formal symbolic representations. They are accessible to users who are not acquainted with formal systems such as HamNoSys or SW. These systems, which rely on observable gesture movements extracted from real-world video, are particularly appropriate for contexts where signers or interpreters lack formal training, and where clarity and accessibility are paramount.

Joksimoski et al. and other studies note that most systems are limited to narrow domains (such as weather or medicine) and do not transfer knowledge well to new areas [81]. The proposed methodology (A1–A4) serves as a roadmap for scientists and developers, helping to consciously choose the architecture of the system. For example, P1 can be used for educational programs in narrow areas, or P3 can be applied to create dynamic web interfaces, if the computing power and data availability of a particular language allow it.

## 5. Conclusions

The comparative analysis of the three paradigms of sign language visualization revealed fundamental differences in pipeline architecture, data requirements, portability between languages, linguistic accuracy, technical complexity, and accessibility.

The analysis demonstrated that paradigms differ fundamentally in the level of formalization and automation. The formal rule-based pipeline (P1) provides fully deterministic gesture generation and a high level of reproducibility, while P2 relies on the interpretation of visual spatial structures, which complicates the transformation into three-dimensional animation. P3 demonstrates the highest degree of automation and is capable of functioning in real time, but its architecture is significantly more complex and less transparent in terms of linguistic interpretation.

Data requirements were identified as a critical factor affecting scalability. P1 provides high linguistic accuracy, but requires considerable expert resources, which limits the speed of expansion of dictionaries and domains. P2 offers a more sustainable growth model by involving the user community but faces technical limitations when moving from 2D to 3D motion. P3, despite minimizing manual linguistic markup, is critically dependent on the presence of large, annotated video datasets, which limit its practical applicability to many low-resource national sign languages.

Formal notations and writing systems ensure high portability of P1 and P2 between languages and subject areas without the need for architectural modification. The portability of P3 is limited by the linguistic specifics of the training data, which can lead to a loss of grammatical adequacy when applying the model to new languages or communicative contexts.

The comparison also highlighted differences related to accuracy, accessibility, and complexity. P1 provides the highest level of linguistic credibility but remains a laborious and niche solution. P2 stands out for its high iconicity and user-centered nature, which increases cultural and social acceptability, but requires further research to improve 2D–3D transformation. P3 demonstrates the most natural and smooth movements but is prone to the loss of non-manual and inflective markers, which are critical for the grammatical completeness of sign languages.

Overall, the findings suggest that no single paradigm can be considered universally optimal, as each reflects different balances between linguistic control, scalability, and accessibility.

In addition to theoretical comparison, the suggested framework provides a pragmatic decision-making instrument for researchers and developers. It enables stakeholders to choose the most suitable paradigm depending on available resources (e.g., data, knowledge, target language) and intended application (e.g., education, media, accessibility tools). This highlights the importance of selecting an approach based on application context and resource availability, particularly for low-resource sign languages.

The conclusions obtained indicate the prospects of hybrid architectures, combining the formal linguistic accuracy of notation approaches with the naturalness of movements and automation inherent in modern AI methods. Such synthesis can become a key direction for the development of future avatar systems and contribute to increasing digital accessibility for speakers of low-resource sign languages.

Subsequent research could experimentally validate the paradigm via case studies and develop and evaluate hybrid pipelines that include symbolic and data-driven elements. Furthermore, broadening the criteria to encompass personalization, affective expressiveness, or user input could enhance its usefulness. Future research may also benefit from integrating linguistic and data-driven representations to improve both interpretability and realism.

**Author Contributions:** Conceptualization, N.A.; methodology, N.A. and M.M.; software, N.A.; validation, N.A., A.Y., E.D. and M.M.; formal analysis, N.A.; investigation, N.A. and N.T.; resources, A.K.; data curation, A.Y., N.A. and N.T.; writing—original draft preparation, N.A.; writing—review and editing, A.Y., M.M., N.A., N.T. and A.K.; visualization, A.K., E.D.; supervision, N.A. and A.Y.;

project administration, N.A.; funding acquisition, A.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR24992875).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

**Conflicts of Interest:** Author Nurzada Amangeldy was employed by the company SignBridge LLP. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. FAQs | World Federation of the Deaf. Available online: https://wfdeaf.org/contact/faqs/ (accessed on 4 January 2026).
2. Deaf History—Europe—Access to Sign Language: A Human Right. Available online: https://deafhistory.eu/index.php/component/zoo/item/human-rights (accessed on 4 January 2026).
3. Six Difficulties We Faced When Thinking About a Sign Language Avatar. Available online: https://innovation.dw.com/articles/six-difficulties-sign-language-avatar (accessed on 4 January 2026).
4. Caselli, N.K.; Sehyr, Z.S.; Cohen-Goldberg, A.M.; Emmorey, K. ASL-LEX: A Lexical Database of American Sign Language. *Behav. Res.* **2017**, *49*, 784–801. [CrossRef]
5. Zou, Y.; Lin, H. A Basic General Service List for Chinese Sign Language. *J. Deaf Stud. Deaf Educ.* **2025**, *30*, 405–418. [CrossRef]
6. Perlman, M.; Little, H.; Thompson, B.; Thompson, R.L. Iconicity in Signed and Spoken Vocabulary: A Comparison Between American Sign Language, British Sign Language, English, and Spanish. *Front. Psychol.* **2018**, *9*, 1433. [CrossRef]
7. Zhao, Y.; Wu, X.; Sun, P.; Chen, H. Relationship Between Vocabulary Knowledge and Reading Comprehension in Deaf and Hard of Hearing Students. *J. Deaf Stud. Deaf Educ.* **2021**, *26*, 546–555. [CrossRef]
8. Convertino, C.; Borgna, G.; Marschark, M.; Durkin, A. Word and World Knowledge Among Deaf Learners With and Without Cochlear Implants. *J. Deaf Stud. Deaf Educ.* **2014**, *19*, 471–483. [CrossRef]
9. Hall, W.C. What You Don't Know Can Hurt You: The Risk of Language Deprivation by Impairing Sign Language Development in Deaf Children. *Matern. Child Health J.* **2017**, *21*, 961–965. [CrossRef] [PubMed]
10. For the Deaf Community, Sign Language Equals Rights | Human Rights Watch. Available online: https://www.hrw.org/news/2022/09/23/deaf-community-sign-language-equals-rights (accessed on 17 January 2026).
11. AI and Machine Translation: A Threat to the Deaf Community?—Deaf Journalism Europe. Available online: https://www.deafjournalism.eu/ai-and-machine-translation-a-threat-to-the-deaf-community/ (accessed on 17 January 2026).
12. BOBSL. Available online: https://www.robots.ox.ac.uk/ vgg/data/bobsl/?utm_source=chatgpt.com (accessed on 5 January 2026).
13. Duarte, A.; Palaskar, S.; Ventura, L.; Ghadiyaram, D.; DeHaan, K.; Metze, F.; Torres, J.; Giro-i-Nieto, X. How2Sign: A Large-Scale Multimodal Dataset for Continuous American Sign Language. In *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Nashville, TN, USA, 2021; pp. 2734–2743.
14. Li, D.; Opazo, C.R.; Yu, X.; Li, H. Word-Level Deep Sign Language Recognition from Video: A New Large-Scale Dataset and Methods Comparison. In *Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*; IEEE: Snowmass Village, CO, USA, 2020; pp. 1448–1458.
15. Belissen, V.; Braffort, A.; Gouiffès, M. Dicta-Sign-LSF-v2: Remake of a Continuous French Sign Language Dialogue Corpus and a First Baseline for Automatic Sign Language Processing. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*; Calzolari, N., Béchet, F., Blache, P., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Isahara, H., Maegaard, B., Mariani, J., et al., Eds.; European Language Resources Association: Marseille, France, 2020; pp. 6040–6048.
16. Camgoz, N.C.; Hadfield, S.; Koller, O.; Ney, H.; Bowden, R. Neural Sign Language Translation. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: Salt Lake City, UT, USA, 2018; pp. 7784–7793.
17. Corpus: British Sign Language Corpus | SL Data Compendium. Available online: https://www.sign-lang.uni-hamburg.de/lr/compendium/corpus/bslcorpus.html (accessed on 17 January 2026).
18. Konrad, R.; Hanke, T.; Langer, G.; Blanck, D.; Bleicken, J.; Hofmann, I.; Jeziorski, O.; König, L.; König, S.; Nishio, R.; et al. *MY DGS—Annotated. Public Corpus of German Sign Language, 3rd Release*; Institute of German Sign Language and Communication of the Deaf (IDGS): Hamburg, Germany, 2020.

19. GitHub—ChelseaGH/KSL-Guide: KSL-Guide: A Large-Scale Korean Sign Language Dataset Including Interrogative Sentences for Guiding the Deaf and Hard-of-Hearing, FG, 2021. Available online: https://github.com/ChelseaGH/KSL-Guide (accessed on 17 January 2026).
20. CSL-Daily Dataset. Available online: https://ustc-slr.github.io/datasets/2021_csl_daily/ (accessed on 5 January 2026).
21. Aziz, M.; Othman, A. Evolution and Trends in Sign Language Avatar Systems: Unveiling a 40-Year Journey via Systematic Review. *Multimodal Technol. Interact.* **2023**, *7*, 97. [CrossRef]
22. Naert, L.; Larboulette, C.; Gibet, S. A Survey on the Animation of Signing Avatars: From Sign Representation to Utterance Synthesis. *Comput. Graph.* **2020**, *92*, 76–98. [CrossRef]
23. Rastgoo, R.; Kiani, K.; Escalera, S.; Athitsos, V.; Sabokrou, M. A Survey on Recent Advances in Sign Language Production. *Expert Syst. Appl.* **2024**, *243*, 122846. [CrossRef]
24. Tuveri, F. A Comprehensive Review of Sign Language Production. *J. Comput. Assist. Linguist. Res.* **2024**, *8*, 1–22. [CrossRef]
25. Bragg, D.; Koller, O.; Bellard, M.; Berke, L.; Boudrealt, P.; Braffort, A.; Caselli, N.; Huenerfauth, M.; Kacorri, H.; Verhoef, T.; et al. *Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective*; Association for Computing Machinery (ACM): New York, NY, USA, 2019.
26. Al-Qurishi, M.; Khalid, T.; Souissi, R. Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues. *IEEE Access* **2021**, *9*, 126917–126951. [CrossRef]
27. Al Abdullah, B.A.; Amoudi, G.A.; Alghamdi, H.S. Advancements in Sign Language Recognition: A Comprehensive Review and Future Prospects. *IEEE Access* **2024**, *12*, 128871–128895. [CrossRef]
28. Torrado Mora, L.D.; Alberto Collazos, C.; Rico Bautista, D. Traducción de Lenguaje de Signos a Texto Mediante Python Con Redes Neuronales LSTM. *RCTA* **2025**, *2*, 150–159. [CrossRef]
29. Wolfe, R.; McDonald, J.C.; Hanke, T.; Ebling, S.; Van Landuyt, D.; Picron, F.; Krausneker, V.; Efthimiou, E.; Fotinea, E.; Braffort, A. Sign Language Avatars: A Question of Representation. *Information* **2022**, *13*, 206. [CrossRef]
30. Izaguirre, E.D.P.; Abásolo, M.J.; Collazos, C.A. BlipBla Mobile and Collaborative App to Teach Lip Reading to Children Who Are Deaf. *IEEE R. Iberoam. Tecnol. Aprendiz.* **2025**, *20*, 182–190. [CrossRef]
31. Naert, L.; Larboulette, C.; Gibet, S. Motion Synthesis and Editing for the Generation of New Sign Language Content: Building New Signs with Phonological Recombination. *Mach. Transl.* **2021**, *35*, 405–430. [CrossRef]
32. Dimou, A.-L.; Papavassiliou, V.; Goulas, T.; Vasilaki, K.; Vacalopoulou, A.; Fotinea, S.-E.; Efthimiou, E. What about Synthetic Signing? A Methodology for Signer Involvement in the Development of Avatar Technology with Generative Capacity. *Front. Commun.* **2022**, *7*, 798644. [CrossRef]
33. Atasoy, M.; Şılbır, L.; Erümit, S.F.; Bahçekapılı, E.; Yıldız, A.; Karal, H. Visual Appearance Features of Sign Language Avatars. *Kastamonu Educ. J.* **2023**, *31*, 386–403. [CrossRef]
34. De Almeida Freitas, F.; Peres, S.M.; De Paula Albuquerque, O.; Fantinato, M. Leveraging Sign Language Processing with Formal SignWriting and Deep Learning Architectures. In *Intelligent Systems*; Naldi, M.C., Bianchi, R.A.C., Eds.; Lecture Notes in Computer Science; Springer Nature: Cham, Switzerland, 2023; Volume 14197, pp. 299–314.
35. Hand Talk: Your Website Accessible in ASL. Available online: https://www.handtalk.me/en/ (accessed on 10 January 2026).
36. Home—Sign Bridge. Available online: https://signbridge.kz/ (accessed on 10 January 2026).
37. JASigning Vhg.2020. Available online: https://vhg.cmp.uea.ac.uk/tech/jas/vhg2020/ (accessed on 10 January 2026).
38. Sevilla, A.F.G.; Lahoz-Bengoechea, J.M.; Conde González, S.; Diaz Esteban, A.; Folgueira Galán, P.; De La Calle Pérez, J. Combining the User Friendliness of SignWriting with the Precision of Linguistic Parameters. In *Adjunct Proceedings of the 25th ACM International Conference on Intelligent Virtual Agents*; ACM: Berlin Germany, 2025; pp. 1–6.
39. Lalitha, S.; Adavi, V. Beyond Words: Speech to Sign Language Interpreter. In *Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*; IEEE: Kamand, India, 2024; pp. 1–6.
40. Varghese, M.; Nambiar, S.K. English To SiGML Conversion For Sign Language Generation. In *Proceedings of the 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET)*; IEEE: Kottayam, India, 2018; pp. 1–6.
41. Aasofwala, N.; Verma, S.; Patel, K. Conversion of Gujarati Alphabet to Gujarati Sign Language Using Synthetic Animation. In *ICT Analysis and Applications*; Fong, S., Dey, N., Joshi, A., Eds.; Lecture Notes in Networks and Systems; Springer Nature: Singapore, 2023; Volume 782, pp. 49–61.
42. Aliwy, H.A.; Ahmed, A.A. Development of Arabic Sign Language Dictionary Using 3D Avatar Technologies. *Indones. J. Electr. Eng. Comput. Sci. (IJEECS)* **2021**, *21*, 609. [CrossRef]
43. Goyal, D.; Goyal, V.; Goyal, L. Automatic Translation of Complex English Sentences to Indian Sign Language Synthetic Video Animations. In *Proceedings of the 17th International Conference on Natural Language Processing (ICON): System Demonstrations*; Goyal, V., Ekbal, A., Eds.; NLP Association of India (NLPAI): Patna, India, 2020; pp. 43–45.
44. Joshi, R.B.; Desale, S.; Gaikwad, H.; Gunje, S.; Londhe, A. A Survey on Sign Language Translation Systems. *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)* **2022**, *10*, 519–525. [CrossRef]

45. Kaur, K.; Kumar, P. HamNoSys to SiGML Conversion System for Sign Language Automation. *Procedia Comput. Sci.* **2016**, *89*, 794–803. [CrossRef]
46. Kumar Attar, R.; Goyal, V.; Goyal, L. State of the Art of Automation in Sign Language: A Systematic Review. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* **2023**, *22*, 1–80. [CrossRef]
47. Jha, A.; Choudhary, K.; Shetty, S. Deep Learning Based Text Translation and Summarization Tool for Hearing Impaired Using Indian Sign Language. In *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods*; SCITEPRESS—Science and Technology Publications: Lisbon, Portugal, 2023; pp. 426–434.
48. Ubieto, V.; Pozo, J.; Valls, E.; Cabrero-Daniel, B.; Blat, J. Sign Language Synthesis: Current Signing Avatar Systems and Representation. In *Sign Language Machine Translation*; Way, A., Leeson, L., Shterionov, D., Eds.; Machine Translation: Technologies and Applications; Springer Nature: Cham, Switzerland, 2024; Volume 5, pp. 247–266.
49. Johnson, R.; Wolfe, R. A Survey of Shading Techniques for Facial Deformations on Sign Language Avatars. In *Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives*; Efthimiou, E., Fotinea, S.-E., Hanke, T., Hochgesang, J.A., Kristoffersen, J., Mesch, J., Eds.; European Language Resources Association (ELRA): Marseille, France, 2020; pp. 107–112.
50. Zerrouki, T.; Slimani, M.F.; Mami, A.; Mazari, R. 3DZSignDB: 3D Avatar SigML Data for Algerian Sign Language. *Data Brief* **2025**, *60*, 111568. [CrossRef] [PubMed]
51. Singh, H.; Mishra, A.; Dubey, R.; Tiwari, V. Generating Avatar Using HamNoSys and SiGML for British Sign Language. In *Proceedings of the 2025 10th International Conference on Signal Processing and Communication (ICSC)*; IEEE: Noida, India, 2025; pp. 427–433.
52. Bouzid, Y.; Khenissi, M.A.; Jemni, M. The Effect of Avatar Technology on Sign Writing Vocabularies Acquisition for Deaf Learners. In *Proceedings of the 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT)*; IEEE: Austin, TX, USA, 2016; pp. 441–445.
53. Barros, R.O.; Vieira, S.Z. The Relationship between Text and Image on Literary Productions in Libras. *Sign Lang. Stud.* **2020**, *20*, 392–410. [CrossRef]
54. Moe, S.Z.; Thu, Y.K.; Nwe, H.M.; Hlaing, H.W.W.; Aung, N.H.; Wai, K.H.; Thant, H.A.; Min, N.W. Development of Natural Language Processing Based Communication and Educational Assisted Systems for the People with Hearing Disability in Myanmar. *Línguas Let.* **2019**, *20*, 32–55. [CrossRef]
55. Bouzid, Y.; Jemni, M. An Avatar Based Approach for Automatically Interpreting a Sign Language Notation. In *Proceedings of the 2013 IEEE 13th International Conference on Advanced Learning Technologies*; IEEE: Beijing, China, 2013; pp. 92–94.
56. Sevilla, A.F.G.; Díaz Esteban, A.; Lahoz-Bengoechea, J.M. Building the VisSE Corpus of Spanish SignWriting. *Lang. Resour. Eval.* **2024**, *58*, 585–607. [CrossRef]
57. Shabina, B.; Sony, S.; Rajeev, R.R. An Approach for Malayalam Sentence to Indian Sign Language Generation Using Synthetic Animations. *AIP Conf. Proc.* **2024**, *3134*, 020002. [CrossRef]
58. Panagi, P.; Yeratziotis, A.; Fotiadis, T.; Mettouris, C.; Papadopoulos, G.A. "Signifier" Video Sharing Platform and Accessible Media Player for Deaf Users. In *Proceedings of the 2024 International Conference on Information Technology for Social Good*; ACM: Bremen, Germany, 2024; pp. 151–157.
59. Myasoedova, M.A.; Myasoedova, Z.P.; Farkhadov, M.P. Multimedia Technologies to Teach Sign Language in a Written Form. In *Proceedings of the 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)*; IEEE: Tashkent, Uzbekistan, 2020; pp. 1–4.
60. Boyes-Braem, P. A Multimedia Bilingual Database for the Lexicon of Swiss German Sign Language. *Sign Lang. Linguist.* **2001**, *4*, 133–143. [CrossRef]
61. Jiang, Z.; Moryossef, A.; Müller, M.; Ebling, S. Machine Translation between Spoken Languages and Signed Languages Represented in SignWriting. In *Proceedings of the Findings of the Association for Computational Linguistics: EACL 2023*; Association for Computational Linguistics: Dubrovnik, Croatia, 2023; pp. 1706–1724.
62. Hoffmann-Dilloway, E. Feeling Your Own (or Someone Else's) Face: Writing Signs from the Expressive Viewpoint. *Lang. Commun.* **2018**, *61*, 88–101. [CrossRef]
63. De Monte, M.T. Writing Sign Languages: In Search of a Definition. *Int. J. Linguist.* **2025**, *17*, 221. [CrossRef]
64. Dhanjal, A.S.; Singh, W. Comparative Analysis of Sign Language Notation Systems for Indian Sign Language. In *Proceedings of the 2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*; IEEE: Gangtok, India, 2019; pp. 1–6.
65. Singh, G.; Goyal, V.; Goyal, L. Notation Systems Applied on Indian Sign Language: A Review. In *Recent Innovations in Computing*; Singh, P.K., Singh, Y., Kolekar, M.H., Kar, A.K., Gonçalves, P.J.S., Eds.; Lecture Notes in Electrical Engineering; Springer: Singapore, 2022; Volume 832, pp. 71–79.

66. Bouzid, Y.; Khenissi, M.A.; Jemni, M. Designing a Game Generator as an Educational Technology for the Deaf Learners. In *Proceedings of the 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA)*; IEEE: Marrakech, Morocco, 2015; pp. 1–6.
67. Ghaziasgar, M.; Bagula, A.; Thron, C. Automatic Sign Language Manual Parameter Recognition (II): Comprehensive System Design. In *Implementations and Applications of Machine Learning*; Subair, S., Thron, C., Eds.; Studies in Computational Intelligence; Springer International Publishing: Cham, Switzerland, 2020; Volume 782, pp. 93–118.
68. Goodwin, A. Signwriting: Ornament as Visual Language—Communicative Decoration. *J. Illus.* **2019**, *6*, 119–136. [CrossRef]
69. Hoffmann-Dilloway, E. Writing the Smile: Language Ideologies in, and through, Sign Language Scripts. *Lang. Commun.* **2011**, *31*, 345–355. [CrossRef]
70. Lobo-Neto, V.C.; Pedrini, H. LSWH100: A Handshape Dataset for Brazilian Sign Language (Libras) Using SignWriting. *Data Brief* **2024**, *56*, 110780. [CrossRef] [PubMed]
71. Maher Ring, D. Typ.Ologies: Reframing Ireland's Vernacular Letterform through the Lens of Heritage. *InfoDesign* **2023**, *20*, 2. [CrossRef]
72. Faisal, M.; Alsulaiman, M.; Mekhtiche, M.; Abdelkader, B.M.; Algabri, M.; Alrayes, T.B.S.; Muhammad, G.; Mathkour, H.; Abdul, W.; Alohali, Y.; et al. Enabling Two-Way Communication of Deaf Using Saudi Sign Language. *IEEE Access* **2023**, *11*, 135423–135434. [CrossRef]
73. Natarajan, B.; Rajalakshmi, E.; Elakkiya, R.; Kotecha, K.; Abraham, A.; Gabralla, L.A.; Subramaniyaswamy, V. Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation. *IEEE Access* **2022**, *10*, 104358–104374. [CrossRef]
74. Dimou, A.-L.; Papavassiliou, V.; Mcdonald, J.; Goulas, T.; Vasilaki, K.; Vacalopoulou, A.; Fotinea, S.-E.; Efthimiou, E.; Wolfe, R. *Signing Avatar Performance Evaluation within the EASIER Project*; European Language Resources Association: Paris, France, 2022.
75. Gajic, D.; Gojic, G.; Dragan, D.; Petrovic, V. Comparative Evaluation of Keypoint Detectors for 3d Digital Avatar Reconstruction. *Facta Univ. Ser. Electron. Energetics* **2020**, *33*, 379–394. [CrossRef]
76. Berrezueta-Guzman, S.; Daya, R.; Wagner, S. Virtual Reality in Sign Language Education: Opportunities, Challenges, and the Road Ahead. *Front. Virtual Real.* **2025**, *6*, 1625910. [CrossRef]
77. Rochow, A.; Schwarz, M.; Schreiber, M.; Behnke, S. VR Facial Animation for Immersive Telepresence Avatars. In *Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*; IEEE: Kyoto, Japan, 2022; pp. 2167–2174.
78. Rochow, A.; Schwarz, M.; Behnke, S. Attention-Based VR Facial Animation with Visual Mouth Camera Guidance for Immersive Telepresence Avatars. In *Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*; IEEE: Detroit, MI, USA, 2023; pp. 1276–1283.
79. DAI—ASLLVD. Available online: https://www.bu.edu/asllrp/av/dai-asllvd.html?utm_source=chatgpt.com (accessed on 5 January 2026).
80. Sign-Language Datasets at Scale: A Comprehensive Survey on Resources, Benchmarks, and Annotation Standards. 30 July 2025. Available online: https://openreview.net/forum?id=B3ruHnLnEw#discussion (accessed on 10 May 2026).
81. Joksimoski, B.; Zdravevski, E.; Lameski, P.; Pires, I.M.; Melero, F.J.; Martinez, T.P.; Garcia, N.M.; Mihajlov, M.; Chorbev, I.; Trajkovik, V. Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities. *IEEE Access* **2022**, *10*, 40979–40998. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.