Tuveri, Franco (2024) A comprehensive review of sign language production. Journal of Computer-Assisted Linguistic Research 8: 1-22. ISSN: 2530-9455.

A Comprehensive Review of Sign Language Production

Franco Tuveri*
CRS4, Center of Advanced Studies, Research and Development in Sardinia, Italy
*Corresponding author: tuveri@crs4.it

Received: 9 January 2024 / Accepted: 15 April 2024 / Published: 15 November 2024

Abstract
Sign languages are made up of phonological, morphological, syntactic and semantic levels of structure that satisfy the same social, cognitive and communicative purposes as other natural languages and represent the most used form of communication between hearing and deaf people. Sign Language Production together with Sign Language Recognition constitute the two parts of this process, as Sign Language Production concerns that part of the communication process that goes from spoken language to its translation into Sign Language, while Sign Language Recognition deals with the recognition of Sign Language. In this article, we want to consider some of the most recent and important studies on Sign Language Production and discuss their limitations, advantages, and future developments.

Keywords: Sign Language, Notation Systems, Machine Translation, Sign Language Production

1. Introduction

Sign languages (SLs) consist of phonological, morphological, syntactic and semantic levels of structure that fulfill the same social, cognitive and communicative purposes as other natural languages. They have evolved naturally, like all languages, but with a structure that uses both manual and non-manual components and possesses mechanisms of evolutionary dynamics and regionalisms. SLs exploit the visual-gestural channel and allow signatory deaf people equal opportunities for access to communication. While spoken languages mainly exploit the vocal-auditory modality, sign languages use the visual-gestural modality, relying on facial expressions, hand movement, body and space used by the signer.

According to the World Health Organization (WHO, 2021), over 5 percent of the world's population, about 430 million people, will require rehabilitation to address their disabling hearing loss (including 34 million children), and about 60 million have severe or complete deafness. It is also expected that, by 2050, nearly 2.5 billion people will have some degree of

DOI: HTTPS://DOI.ORG/10.4995/JCLR.2024.20983
PUBLISHER: UNIVERSITAT POLITÈCNICA DE VALÈNCIA.

hearing loss, and at least 700 million will require hearing rehabilitation. In Italy, there are 7 million people with hearing impairment, corresponding to 12.1 percent of the population, a portion of whom are completely deaf. While in the world SLs are used by about 70 million signers, there are about 100,000 signers in Italy, of whom about 40,000 are deaf. Many deaf people, in fact, do not know and do not use SL but communicate through spoken language (oralists) and/or by lip-reading.

In a dialogue between a hearing person and a deaf person, it is possible to observe a number of difficulties and problems, some of which are immediately visible while others not so much. The two people communicate in different languages. The first essentially uses vocal language, sometimes underlining with facial expressions, gestures and even body movements the passages on which s/he wants to express greater emphasis with co-expressive movements. The deaf person instead communicates with a gestural language that is expressed through the movement of the hands, non-manual signals as well as with facial expressions and body movements. In the first case, the movements are only co-expressive, while in the other, the body movements have syntactic and semantic value and are part of the linguistic convention that represents Sign Language. Communication could also take place through writing as long as the deaf person knows how to write, but unfortunately this is not a condition that all deaf people have in common. There is, therefore, a need for tools that can bridge existing communication differences (Marchetti et al., 2012).

Sign Language Production (SLP), together with Sign Language Recognition (SLR), are the two necessary parts of the mentioned communication process. Specifically, SLP concerns that part of the communication process that goes from spoken language to its translation into Sign Language, while SLR deals with the recognition of Sign Language.

The review analyses, in a non-exhaustive way, the work done in the field of SLP research according to different technologies, trying to gather the most important studies of the last years. Its primary goal is to prioritize studies that have most recently investigated the specific aspects that have been identified as topics of interest and as building blocks for the implementation of an SLP system. A large number of articles and reviews with SLP as a central theme have been collected through searches in the main journals and conferences. Furthermore, only papers written in English were included in this study.

The article aims to provide a comprehensive review of the literature regarding the current state of research on SLP, focusing on the following research questions:

- What are the aspects and technologies that converge in SLP?
- What are the main notation systems?
- What are the critical issues related to the Sign Language Machine Translation (SLMT)?
- What are the technologies used in SLMT?
  - What are the methods and measures in NMT?
  - What are the methods in the RBMT?
- What are the main dedicated and multilingual datasets?

The article, therefore, identifies the critical issues related to the components that compose the

SLP process by reviewing the main notation systems, the main technologies and methodologies used in SLMT, and the main datasets available. The most recent studies are then highlighted, taking into account the different approaches used and the advantages and disadvantages associated with the implementation choices made by the research groups.

The remainder of this paper is organised as follows. Section 2 presents an introduction to Sign Language Production and summarises the main related concepts. Section 3 focuses on the most used notation systems. Section 4 introduces the Sign Language Machine translation techniques. Section 5 reviews the available datasets. Finally, section 6 provides the conclusions.

2. Sign Language Production

SLP, the automatic translation from spoken language sentences into sign language sequences, is a challenging problem that involves both linguistic and visual aspects (Rastgoo et al., 2021). It requires understanding the grammar and vocabulary of both the spoken and the sign languages, as well as generating realistic and expressive sign motions. The challenges regarding SLP are an expression of the many difficulties related to the different ways in which translation from speech to sign language is achieved (Rastgoo et al., 2022).

The transcription of communication in sign language is one of the most challenging tasks because it involves the simultaneous use of manual and non-manual information, gestures in spatial motion, and a combination of conventional signs and gestures (Slobin et al., 2001). The components of gestures can be summarised as given below:

- Manual Gestures: Shape, Movement, Location, Orientation
- Non-Manual Gestures: Facial Expression, Head tilting, Mouthing, Shoulder raising.

Another critical issue is the ability to reproduce a sign starting from a text or speech in the real world. In an ideal situation, respect for the grammatical forms and syntax of the source language strongly influences the performance of the systems responsible for translation into SL. Translation turns out to be a complex problem as it cannot occur as a simple mapping of the individual words that make up a vocal/textual message into the corresponding sign. In fact, there are differences between tokenization and word ordering in spoken and signed language (Rastgoo et al., 2021). It is first necessary to identify the linguistic features specific to SLs that researchers need to consider during their modelling. Moryossef and Goldberg (2021) identify these features as: Phonology, Simultaneity, Referencing, and Fingerspelling.

Phonology is the study of the smallest part of the language that produces meaning. In a spoken language, like Italian or English, a phoneme is a unit, a sound that expresses meaning. If we replace one vowel with another in a term, the meaning of that term will change to something completely different, e.g. rice and race.

The same happens in a similar way in Sign Languages, where there are five parameters or characteristics, which, if realized in different ways, can change the meaning of a sign. Each of the five Sign Language parameters must be used accurately for a sign to be correct. The five parameters are: hand shape, palm orientation, location, movement, and non-manual signs.

Accurately using the five parameters affects meaning and impacts the understanding of SLs.

Slonimska et al. (2020) stated that signers use simultaneous and iconic constructions in their sign language in order to increase communicative efficiency on occasions when they are asked to encode information-rich events. Bellugi and Fischer (1972) asserted that an SL sign takes about twice as long to produce as an English word despite the speed of information transmission between the two languages being similar. This result is explained through the mechanism of simultaneity, which allows SLs to compensate for the slower rate of sign production. In fact, SLs can simultaneously use multiple visual signals to convey information (Sandler, 2012). The signer can, in fact, produce the sign of an object with one hand and simultaneously indicate the real object. Another way of expressing simultaneity can occur, similarly in spoken languages, using different tones of voice or through facial expressions or the position of the torso. In the same way, the signer can enrich the signed sentences with affective information (Liddell, 2003; Johnston and Schembri., 2007) through facial expressions and, in this way, modify the meaning of adjectives, adverbs, and verbs.

Hand movements are complementary gestures that, in general, are not part of spoken languages but are often co-expressive of meaning (McNeil, 1992). In the case of SLs, hand movements are called signs and are part of the linguistic convention that represents Sign Language.

In referencing, the signer can introduce referents into their speech by indicating their location in space or by assigning a region of sign space to a referent not present and pointing to this region to refer to it (Rathman & Mathur, 2011; Schembri et al., 2018). Although this can also occur in oral speech, in SLs these elements, which can vary in meaning depending on the conformation of the hand, have syntactic and semantic value (Kendon, 2004; Kooji et al., 2006; Pfau, 2011).

Signers can also establish relationships between multiple referents in sign space by using pointing or by taking on the role of referents by using body movement or gaze. In morphology, referencing also occurs when the directionality of a verb goes from the position of the subject towards the position of its object (Fenlon et al., 2018).

Referencing also intervenes in the use of classifiers or depicting signs (Supalla, 1983), which serve to describe the referent. Classifiers are complex structures that consist of a specific manual configuration associated with a movement. Classifiers can be used to represent how the referent relates to other entities. They are signs usually represented with a single hand that denote animate or inanimate referents in their entirety, considering the form or semantic category to which they belong.

Fingerspelling is a method of spelling words using hand movements and consists of a set of hand gestures used to represent the alphabet or numbers, but it is also found in many sign languages to spell out names of people and places for which there is no sign. It has often become integrated into the signed languages as another linguistic strategy (Padden, 1998; Montemurro and Brentari, 2018).

Patrie and Johnson (2011) identified three different forms of fingerspelling in American Sign Language (ASL):

- Careful or slower spelling, in which each letter is clearly marked.
- Rapid or quick spelling, in which letters are often incomplete or contain remnants of other letters in the marked word.
- Lexicalized fingerspelling, or the way in which a sign is represented by at most two letters of the word. Lexicalized fingerspelled signs are tipically indicated by the symbol # preceding the sign (Battison, 1978). For example, fix is lexicalized as #FX, yes as #YS, bank as #BK, etc.

3. Notation Systems

Writing is a common element in practically all spoken languages. Sign languages, on the other hand, based on a gestural-visual channel, require a different form of representation, as writing, due to the monolinearity of speech production, does not allow the multilinearity of SL to be easily reproduced.

To date, there are no writing systems that can represent multilinearity since all transcription systems are designed for the monolinearity of spoken languages. To give an example, there is a difference between suggesting a melody by reproducing the sequence of notes that compose it and playing the complete score by the orchestra of all the instruments. Although the sequence of notes may suggest the musical motif, the symphony played by the orchestra will be considerably more complex and rich.

Some crucial aspects in the creation of an SLP system highlight the importance of the use of notations. From this perspective, the most easily identifiable critical points are the following::

1. Loss of information from the point of view of phonology, both in the source language, in speech to text, and in the translation into sign languages (Camgoz et al., 2020; Yin et al., 2021). This occurs because a series of phonological elements of the SLs can be lost in transcription;
2. correct identification of the sign to represent;
3. information representation: despite the possibility of using increasingly refined notation systems (GLOSS, HamNoSys with SiGML, SignWriting, etc.) for the passage of the transcribed message into SL, there are pros and cons for each of the existing approaches. Currently, the main approaches (Rastgoo et al., 2021) are as follows:
    a) Avatar;
    b) NMT;
    c) Motion Graph;
    d) Pose-to-Video (Moryossef et al., 2023).

In the collected studies, the authors' need to define accessible and user-friendly specific tools to make writing glyphs easier and faster emerges in several cases. In this sense, Doan et al. (2019) defined a virtual keyboard-based tool, which ideally allows the automatic generation of specific glyphs using key combinations or glyph selection via a specific GUI (Graphical User Interface). Other researchers have instead proposed notation systems based on ELAN (Cormier et al., 2015; Mukushev et al., 2022) or similar tools. Although these tools are used for the purpose of training

NMT algorithms for sign recognition, they refer to pre-existing notation systems and allow the creation of parallel corpora in the various notation systems. As previously said, the aspect of notation is considered from the point of view of translation from spoken (or written) language to SL and not vice versa.

Other researchers have proposed various solutions over time based on glosses or type font systems (Johnston, 2008a, 2008b). Both solutions have pros and cons although, at the moment, two typographic systems are currently the most used: HamNoSys and SignWriting, to which specific paragraphs are dedicated later.

Of the various notation systems, only a few of the most recent and mainly used ones will be considered:

1. Stokoe (For historical and preparatory reasons for HamNoSys.);
2. HamNoSys and SiGML (Signing Gesture Markup Language);
3. SignWriting and SWML (SignWriting Markup Language);
4. SL-GLOSS;
5. ID-GLOSS;
6. The typographic system Typannot.

Notation System	**Sign Language Dependant**	**Non-Manual Features Support**	**Objective**	**Arrangement**	**Computer**
Stokoe	Yes	No	Dictionary or Academic	Linear	Custom Font or ASCII codes
HamNoSys	No	Yes	Academic	Linear	Custom Font or ASCII codes
SignWriting	No	Some	Public Use	Pictorial	ASCII or UNICODE
(SL and ID) GLOSS	Yes	Yes	Academic	Linear	Custom Font or ASCII codes
Typannot	Yes	Movement and Location[1]	-	Linear and Pictorial	Requires a dedicated keyboard

TABLE 1. COMPARISON OF THE MAINLY USED AND RECENT NOTATIONS

In Table 1, the notation system comparison, except for Typannot, is provided by Khan et al. (2021).

3.1. Stokoe

The notation was presented in 1960 in the form of a publication by William Stokoe (1960), for writing ASL (American Sign Language) and later for British SL and Australian Aboriginal SL. It

[1] Location is the place where the hand is placed at the beginning of a sign.

represents the first phonetic script used for Sign Language and has decisively influenced subsequent notations defined in this area. Stokoe notation is defined linearly from left to right and can be written using a writing program that has the appropriate fonts installed. This notation is based on a set of symbols that makes it possible to represent the four fundamental parameters (considered as such by Stokoe) of the Sign Languages (Stokoe, 1960; Stokoe et al., 1965), namely: the Hand Shape (HS), the orientation of the hands, the place where the sign is made, and the set of hand movements to perform the sign. In contrast, non-hand signs, such as body postures, gaze and facial expressions, are not taken into account (Cuxac, 2000; Pizzuto et al., 2006). Stokoe's approach, like the Hamburg Notation System, by focusing exclusively on the gestural components of the hands, cannot fully express the marked speech nor, conversely, allow for a faithful reading of the transcribed sign. Figure 1 shows an example of how the ASL sign THANK-YOU could be notated in the Stokoe notation (Anderson et al., 2022, 143).

FIGURE 1. A POSSIBLE STOKOE NOTATION FOR "THANK-YOU" (ASL)

3.2. HamNoSys and SiGML

The Hamburg Notation System (Prillwitz, 1989), or HamNoSys, was developed in 1985 at the University of Hamburg, Germany. HamNoSys at his latest version (Hanke, 2004) filled some minor gaps and introduced some shortcuts, but has also tackled significant issues related to using HamNoSys in sign language generation. In order to enhance its utility for this purpose, it has been enriched with additional systems enabling the encoding of non-manual behaviours with a level of detail previously unattainable in HamNoSys.

HamNoSys is capable of describing all signs used in all sign languages. It does not rely on sign language conventions that differ from country to country and thus can be used internationally. The HamNoSys notation for a single sign describes the initial hand posture through several features: the shape of the hand, the orientation and position of the hand, as well as actions that modify this posture sequentially or in parallel. For two-handed signs, a symmetry operator precedes the initial posture notation. The operator indicates how the description of the dominant hand is reflected in the position of the non-dominant hand when a different action is not otherwise specified (Hanke, 2004). Like other systems, it is meant to be used in linguistics and includes details such as phonemes that lead to long, complex segments.

Although HamNoSys symbols are readable by humans, it is highly preferable to use notation systems that can provide intermediate representations of the language via XML-based markup systems that are computationally compatible and readable (Zwitserlood et al., 2005). For instance, Signing Gesture Markup Language (SiGML) describes HamNoSys symbols in XML tags form. SiGML (Zwitserlood et al., 2005) representation made from HamNoSys notation of sign language is readable by 3D rendering software (Neves et al., 2020). The article proposes an open-source, language-independent library that allows real-time conversion from HamNoSys to SiGML. The library can be included in Unity or other similar 3D development platforms.

Like Stokoe's notation, HamNoSys presents some limitations as highlighted by some research groups (Fabbretti and Pizzuto, 2000; Bergman et al., 2001; Pennacchi et al., 2001), among which in particular it is possible to highlight: complexity in both writing and reading; lack of connection with the actual form of the sign; the frequent difficulty of transposing speeches into signs; the lack of important parts in the phonology of the sign (e.g. the non-manual components); the memorization of the criteria used to transpose the signs; the difficulty in reusing data; the dynamism of the sign and the correct characterization of the sign space. In Figure 2, (Anderson et al., 2022, 144), the SignWriting and HamNoSys symbols represent the handshape for the letter U in ASL.

FIGURE 2. SIGNWRITING (LEFT) AND HAMNOSYS (CENTER) REPRESENTATION OF THE "U" HANDSHAPE

3.3. SignWriting and SignWriting Markup Language

Sign Writing (SW) was created by Valerie Sutton (Sutton & Gleaves, 1995), a former choreographer who had previously devised a similar system for describing dance movements. SW is proposed as a "hybrid" representation system that allows signs to be codified through symbols placed on a two-dimensional plane (such as drawings). It has been used in some works with the aim of overcoming numerous problems found in other notation systems, facilitating the work of writing and transcribing the marked speeches (Gianfreda et al., 2009). SW is based on a set of "glyphs" that can be combined with each other within a two-dimensional space.

According to Lucioli et al. (2008), a particularity of SW is the ease with which deaf people with good proficiency in LIS, the Italian Sign Language, are able to learn both to write and read it, although at first glance due to its 35,534 glyphs it might seem excessively complicated and not suitable for writing/transcribing the signs (Valeri, 2003).

Again according to Gianfreda et al. (2009), SW would allow, compared to other notation systems, to take into account the manual and non-manual components of the sign, the dynamism of the sign and, above all, the typical structures of the SL complex structures, formally distinct from signs defined as "standard" on the basis of semiotic criteria and specific articulatory traits, such as "classifiers" and impersonation mechanisms (Cuxac, 2000). Thanks to the choice of vertical writing, with SW it is then possible to highlight some structural characteristics of the SL, such as spatial relationships, directionality and movement, which convey fundamental linguistic information.

SWML, the SignWriting Markup Language (da Rocha Costa and Dimuro, 2001), is an XML-based format for the storage and processing of sign language documents written in the SignWriting system and for insertion of sign language texts in HTML documents. The DTD representation of the SWML format is available online.[2] It should be emphasized that SWML encoding of SignWriting is concerned only with the graphical features of texts but in no way diminishes the importance of the semantic aspects of SLP. In fact, SignWriting is not about describing the meaning of signs but the gestures that compose them.

This, according to the authors, is why no linguistic concepts had to be used to structure the representation of signs in SWML but only the concepts related to the graphical construction of SignWriting texts.

A note should be added: the SWML encoding of SignWriting texts considers only the graphical features of such texts. This does not dismiss the importance of semantic issues in sign language processing.

3.4. SL-GLOSS

Glosses are the most commonly used written form to annotate a text in Sign Language (Angelova et al., 2022), according to the syntax of SL, where each sign corresponds to the transcription of a written gloss. Glosses are mono-linear labels, as is writing for verbal languages; however, a major limitation of the use of this notation is represented by the fact that it is not possible to completely represent all the information expressed. As a result, there is a loss of information at the semantic level (Camgoz et al., 2020; Yin et al., 2021). This occurs because a series of phonological elements of the SL are lost in the transcription, such as posture, head movement, facial expressions and expressions made in parallel.

It is easy to imagine that this method introduces several problems. Translating a spoken sentence into sign glosses is a nontrivial task, as is the order and number of glosses that do not correspond to the words of the sentence in the spoken language. Moreover, treating sign language as a concatenation of isolated glosses loses the context of the sentence and the meaning conveyed by the non-manual discursive elements. Although there are several works on gloss-to-text translation in SL, a notable attempt has been made by Moryossef et al. (2021) as part of a translation in a context of scarcity of available resources, in which gloss-text pairs and their linguistic properties are explored in detail. Translation is done with rule-based heuristics for SL gloss generation. Gloss notation is written using capital letters. Table 2 (Amin et al., 2021) shows, as example, some pairs of English sentences represented in ASL.

English sentences	ASL GLOSS
What is your name?	NAME YOU WHAT ^{WH}

[2] SWML: https://www.signwriting.org/forums/software/swml/swml06.html (Accessed on 4 Dec. 2023)

English sentences	ASL GLOSS
He doesn't like pizza.	PIZZA IX-boy DOESN'T-LIKE
Help me.	HELP-ME
Today is Friday, October 28th.	NOW+DAY FRIDAY fs-OCT 28

TABLE 2. ASL GLOSS REPRESENTATION OF SOME ENGLISH SENTENCES

3.5. ID-GLOSS

According to Johnston (2008a), as important as it is to develop a transcribed corpus for SL, it is even more important that it be annotated and that signs be identified using unique annotations based on glosses. The definition of ID-GLOSS versus GLOSS is that ID-GLOSSES are not translations of signs but labels: "*An ID-gloss is the (English) word that is used to label a sign all of the time within the corpus, regardless of what a particular sign may mean in a particular context or whether it has been systematically modified in that context*" (Johnston 2008b, 84).

By consulting the Auslan lexical database, it is possible to find the ID-gloss standard of the signs. A public version of the database is provided by the ASL SignBank.[3] Through the Auslan lexical database, it is possible to associate a sign with an existing ID-GLOSS or create a new one.

3.6. The Typographic System Typannot

The Typannot typographic system (Doan et al., 2019) is a transcription system inspired by the most widespread systems today, HamNoSys, developed for research purposes, and SignWriting, generally considered more suitable for handwriting and for the ability to convey the meaning of signs in the most effective way.

In the Typannot structure, each sign of the SL has three different levels of information: the parametric level, the part level and the feature or feature combination level. The three levels define the sign and, therefore, its transcription.

This structure is repeated for the following parameters of the SL: handshape, location, movement, and facial expression. The visualisation at a typographic level occurs following the four design principles indicated as genericity, readability, modularity and scriptability.

The system allows the transcription of an SL sign using two forms of representation, one generic and one composite, which serve to facilitate transcription using typographical characters of the system defined with the aim of "*allow the easy writing (scriptability) of concatenated information (genericity), corresponding to low-level features (modularity) into a highly readable glyph (readability)*" (Doan et al., 2019, 18).

A virtual keyboard is required and the system is able to manage an input level that works on specific types of user interfaces that operate at the level of parameters, gestures and semi-

[3] ASL Signbank: https://aslsignbank.haskins.yale.edu/ (Accessed on 4 Dec. 2023)

compound characters in order to make transcription easier. Work has been completed on HS and is currently under development on other aspects of facial expressions and transcription interfaces via the Typannot Keyboard.

4. Sign Language Machine Translation

Notation systems constitute a form of representation of dictionaries on which it is possible to build Machine Translation systems for SL. The transition from text to any form of notation effectively highlights the linguistic differences that exist between spoken and signed languages. The monolinear characteristic typical of the former must take a multilinear form for a correct representation of the message to be translated. This aspect is important in order to limit the loss of information that we have already talked about but, unfortunately, the lack of parallel corpora large enough to train a neural machine translation (NMT) system must also be considered, with the exception of corpora based on glosses.

A further critical aspect in the context of sign language translation systems is the lack of a reference base that can be considered reproducible and reliable, without which it is extremely difficult to measure the progress and effectiveness of new methods and systems. The translation system proposed by Moryossef et al. (2023), extending the work of Stoll et al. (2018; 2020), attempts to fill this gap through an open-source system that implements an approach that sequences text-to-gloss-to-pose-to-video steps. In their work, three alternatives are considered in the transition from text to gloss: the use of a lemmatizer, a rule-based translation component that reorders and eliminates terms, and an NMT system. In the transition from gloss-to-pose, data relating to the three sign languages are used: Swiss German Sign Language (DSGS), Swiss French Sign Language (LSF-CH), and Swiss Italian Sign Language (LIS-CH). The sequence involves a sentence in a vocal language being translated, for example, according to a rule-based method, into a sequence of glosses. The glosses are used to search for corresponding videos, coming from a lexicon in one of the languages present (e.g. DSGS). Poses are then extracted from these videos and concatenated to create the sequence that defines the sentence, which is transformed again into a video using the pose-to-video model.

The text-to-gloss phase, rightly contested by some research groups (Yin and Read, 2020; Muller et al., 2022) due to the fact that glosses are unable to effectively represent sign language, can be, however, a good way to improve the level of communications between hearing and deaf people. Proof of this is the existence of numerous works that follow this approach.

Several working groups have used advanced deep learning techniques such as Recurrent Neural Network (RNN) with Luong attention (Stoll et al., 2018; Stoll et al., 2020), Gated Recurrent Unit (GRU) (Amin et al., 2021), and Transformers (Saunders et al., 2020; Ouargani and Khattabi, 2023). The performance due to these techniques guarantees that the translation in the form of GLOSS accurately represents linguistic nuances and has the correct quality requirements.

The performance of the above methods was compared in Table 3 (Nuňez-Marcos et al., 2023; Ouargani and Khattabi, 2023) on the RWTH-PHOENIX-Weather 2014T (PHOENIX 14T) (Camgoz et al., 2018) text-GLOSS parallel corpus. The Ouargani and Kettaby (2023) method outperformed both the GRU and the RNN with Luong attention and performed higher scores of the method used by Saunders et al. (2020) on ROUGE and BLEU1.

Methods	**ROUGE**	**BLEU1**	**BLEU2**	**BLEU3**	**BLEU4**
RNN with Luong attention (Stoll et al., 2020)	48.10	50.67	32.25	21.54	15.25
Transformers (Saunders et al., 2020)	54.55	55.18	**37.10**	**26.24**	**19.10**
GRU (Amin et al., 2021))	42.96	43.90	26.33	16.16	10.42
Transformers (Ouargani and Khattabi, 2023)	**55.18**	**63.60**	28.50	15.20	9.00

TABLE 3. ROUGE AND BLEU SCORES COMPARISON

BLEU (BiLingual Evaluation Understudy) (Papineni et al., 2002) and ROUGE (Recall-Oriented Understudy for Gisting Evaluation) (Lin, 2004) scores are metrics commonly used to assess the quality of machine-generated translations.

In parallel to NMT systems, there are examples of Rule-Based Machine Translation (RBMT) systems applied to text-to-gloss. RBMTs represent a class of methods linked to a set of rules, typically created by linguists.

In the direct approach, the process proceeds with a simple word-for-word translation, without any deep analysis or generation. The transfer-based method consists of analyzing, syntactically and semantically, the sentence structure and, based on the word-to-word translation, generating the text in the target language according to the linguistic rules of the target language. In the interlingua method, the source language is analysed and converted into an abstract representation independent of the language, called interlingua. On this basis, the source language is converted into the target language. There are thus two phases of the translation process: from the source language to the intermediate language and from the intermediate language to the target language. The transfer approach requires knowing the source and target languages, and is therefore language-dependent. The interlingua approach carries out a more in-depth analysis of the structure of the sentences of the source language and also provides information of a semantic nature. Both methods naturally have advantages and disadvantages. Figure 3, inspired by the Vauquois Triangle (Vaquois, 1968), provides more details on each of these approaches.

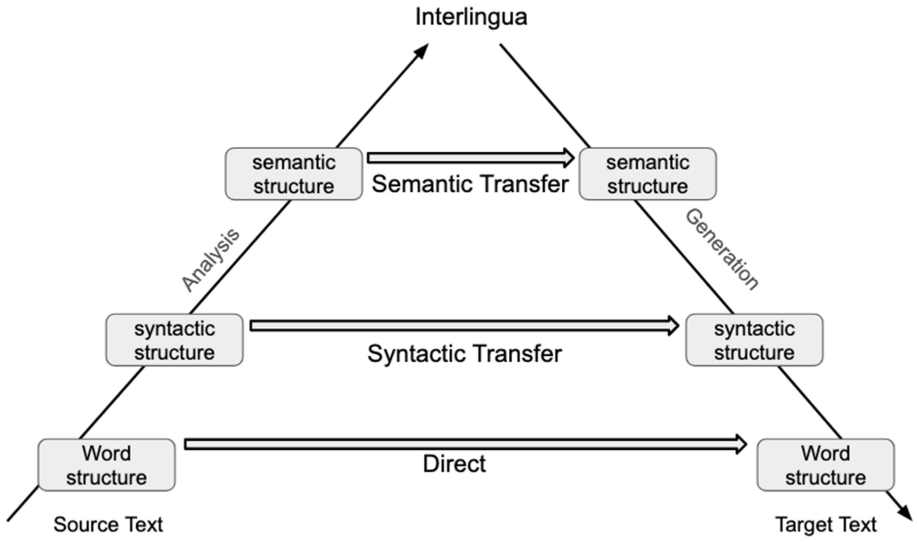

FIGURE 3. THE VAUQUOIS TRIANGLE OF TRANSLATION

The availability of resources often pushes researchers to make precise choices in their approach to translation systems. This is the case of the work carried out by Kouremenos et al. (2018). Due to the limited availability of data for Greek Sign Language (GLS), they created a system based on the RBMT approach. An analysis of the text is first carried out through POS-tagging and chunk-tagging. The Transfer module incorporates a bilingual lexicon and specific knowledge from the rules database, "to transfer the Greek constituency tree structure into the corresponding GSL constituency tree structure" (Kouremenos et al., 2018). The sorting of the glosses is then carried out.

Another study (Ghosh and Mamidi, 2022) was carried out from English to Indian Sign Language, introducing the identification of multi-word expressions and management into the transfer process, with possible replacement of synonyms to produce translations of greater semantic quality. A previous study (Porta et al., 2014) concerns a translation system between written Spanish and glosses according to LSE, Spanish Sign Language. Also, in this case, the approach followed is the RBMT one. The study places emphasis on the syntactic structures of Spanish and on lexical-semantic relationships. Particular attention was paid to classifiers, important constructs of sign languages, in order to make them translatable by identifying Spanish prepositional expressions having locative and temporal meanings.

A characteristic that certainly identifies RBMT systems dedicated to text-to-gloss translation in sign languages is linked to the scarcity of resources available to apply different Machine Translation techniques. With some exceptions, most SLs in the world are not supported by sufficient data structures, datasets, and parallel corpora which are suitable for the application of NMT techniques.

5. Dedicated cross-lingual datasets

The datasets related to SLP are often built during the SL Recognition process, which is the process of recognizing a Sign and its transcription. While SL Recognition benefits from the availability of several large-scale and annotated datasets, public access to similarly extensive datasets for SLP remains limited.

As with modern language corpora, sign language corpora should be representative and well-documented, including relevant metadata and machine readability, facilitating consistent and systematic annotation and tagging (Johnston, 2008a). Achieving this goal requires specialised technology such as ELAN, adherence to standards and protocols exemplified by IMDI metadata descriptors,[4] and the use of transparent and universally accepted grammatical tags such as grammatical class labels (Crasborn et al., 2007).

In producing the datasets used for the recognition of SL signs, several research groups have worked on videos of television programs, such as the RWTH-PHOENIX-Weather corpus (Forster et al., 2012), scanned by professional interpreters who describe the information expressed in the program in sign language. The annotation was performed using ELAN and consists of gloss sentences with words and sentence boundaries and a manual translation of the glosses into written German. In the RWTH-PHOENIX-Weather 2014T dataset (Camgoz et al., 2018), the sign languages used are German (DSG) and American (ASL). As in many other cases, the datasets are linked to specific national SL and, therefore, unusable for recognition systems created for different languages.

As mentioned previously, gestures, the signs used in sign languages, include facial expressions, the localization of the gesture in space, and the movement of the hand, arm and body, and this makes gesture recognition very complicated. Gestures are often captured with useful tools for recording and which are used to optimally create the videos on which the sign recognition tools operate: gloves equipped with sensors and wired or particularly complex tools such as the Panoptic studio used in the case of the creation of the How2Sign dataset (Duarte et al., 2021). The Panoptic studio is a dome with over 500 cameras pointing inside and making it possible to create very precise 3D reconstructions of moving objects and people.

Given its complexity, gesture acquisition necessarily requires compromises between acquisition times, measurement accuracy, production accuracy and spontaneity. Furthermore, it is necessary to calibrate the instruments used and synchronize the data taken through the various channels; the acquisition must comply with precise rules regarding the backgrounds, the distance of the signer from the camera, the ambient brightness, precautions in not covering the hands between of them during the execution of the sign and many other small details.

Some datasets developed in recent years (Rastgoo et al., 2021) are listed below:

1. LinguaSign was an international project funded by the European Union. In its Facebook page,[5] it is possible to find various information regarding the ongoing activities of the working group. It is also important because of its "Catalogue with Signing Avatar",

[4] https://archive.mpi.nl/forums/t/imdi-metadata-information/2933 (Accessed on 26 Mar, 2024)
[5] LinguaSign Primary School Interactive Language Resources: https://www.facebook.com/LinguaSign/ (Accessed on 4 Dec. 2023)

available online.[6]

2. RWTH-PHOENIX- Weather 2014T (Camgoz et al., 2018) is one of the most used datasets in Sign Language Translation. The RWTH-PHOENIX-Weather dataset is built on videos of weather broadcasts and provides spoken language translations and gloss-level annotations in German Sign Language. Nine different signers appear in the videos, the vocabulary contains 1,066 different signs, and the translations into the German spoken language contain a vocabulary of 2,887 different words. The corpus was created with the help of professional sign language interpreters and was annotated using sign glosses by deaf specialists. It was automatically transcribed, manually verified, and normalised.
3. How2Sign (Duarte et al., 2021) consists of a parallel corpus of over 80 hours of sign language video with sentence-level alignment for more than 35k sentences and is defined by the authors as "*a large-scale multimodal and multiview continuous American Sign Language dataset*" (Duarte et al., 2021, 2). Videos have been recorded with multiple RGB and a depth sensor and corresponding gloss annotations. A three-hour subset was further recorded in the Panoptic studio with more than 500 cameras that enabled high-quality 3D keypoints estimation. The authors evaluated the real-world impact of How2Sign by verifying with ASL signers whether the videos, made from the dataset, are indeed understandable.

6. Conclusions and Future Directions

Sign language production is fundamental in the process of translating from spoken and/or written language to sign language. This article has highlighted, in the various steps that make up the SLP, the results of the most significant studies carried out in recent years in the search for solutions that can bridge the communication gap that still exists today between the deaf and hearing communities. The need to use a notation method that is able to best represent the different parameters that characterise the SL was underlined, thus guaranteeing the reproducibility and written memory of what was signed in the Sign Language. The difficulties in this area derive from the fact that the visual-textual channel on which LP is based requires, in transcription, a multilinear representation which is absent in the writing of spoken languages.

An analysis of recent literature was carried out regarding in particular the main parts that make up the translation process from spoken language to SL. In this process, the most commonly used notations that allow us to define a written form of sign language were taken into consideration. Among the various notations, the currently most used in research, both for RBMT and NMT systems, is the one based on glosses, despite its numerous limitations. This is mainly due to the lack of adequate datasets for bilingual translations to and from sign languages. On the other hand, Zhou et al. (2021) defined a model that does not use glosses and demonstrated that, given sufficiently large datasets, NMT systems would outperform gloss-based systems. Some of the main datasets relating to SLPs, the construction methods and the

[6] LinguaSign Catalogue with Signing Avatar: https://vhg.cmp.uea.ac.uk/demo/LinguaSign/ (Accessed on 4 Dec. 2023)

implementation difficulties due to the complexity of the acquisition techniques and tools were then presented and described.

The lack of data sets is a highly limiting factor for NMT models and currently represents one of the major challenges for the construction of SLP systems. The critical issues described in the article are also reflected in the daily lives of deaf people through the lack of useful tools and applications for communicating with other people in society. Furthermore, many applications normally available on the Internet are not currently usable by deaf people and, although efforts are being made to create tools useful for building communication bridges between the deaf and hearing communities, we are still far from being able to have tools that are not limited to specific scenarios.

In conclusion, while the progress made by the most recent deep learning models and the great work done to extend the domain of existing datasets represent an important result for the advancement of SLP, it is also true that further efforts are needed to provide tools capable of compensating for the existing deficiencies in many areas, sometimes even of minor importance, but not yet filled due to the lack of qualitatively adequate tools.

References

Amin, Mohamed, Hesahm Hefny, and Ammar Mohammed. 2021. "Sign Language Gloss Translation using Deep Learning Models". *International Journal of Advanced Computer Science and Applications* 12(11): 686-692. doi: 10.14569/IJACSA.2021.0121178.

Anderson, Catherine, Bronwyn Bjorkman, Derek Denis, Julianne Doner, Margaret Grant, Nathan Sanders, and Ai Taniguchi. 2022. *Essentials of Linguistics, 2nd edition*. eCampusOntario. Ebook ISBN: 978-1-927565-50-6. https://ecampusontario.pressbooks.pub/essentialsoflinguistics2/ (Accessed on Mar. 27, 2024).

Angelova, Galina, Eleftherios Avramidis, and Sebastian Möller. 2022. "Using Neural Machine Translation Methods for Sign Language Translation". In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* 273–284. doi: 10.18653/v1/2022.acl-srw.21.

Battison, Robbin. 1978. *Lexical borrowing in American Sign Language*. Linstok Press. Silver Spring, Maryland.

Bellugi, Ursula, and Susan D. Fischer. 1972. "A comparison of sign language and spoken language". *Cognition* 1: 173-200.

Bergman, Brita, Penny Boyes-Braem, Thomas Hanke, and Elena Pizzuto. 2001. "Sign transcription and database storage of sign information". Special Volume, in *Sign Language and Linguistics* 4(1/2): 3-7. doi: 10.1075/sll.4.12.02ber.

Camgoz, N. Cihan, Simon Hadfield, Oscar Koller, Hermann Ney, and Richard Bowden. 2018. "Neural Sign Language Translation". In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, UT.

Camgoz, N. Cihan, Oscar Koller, Simon Hadfield, and Richard Bowden. 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 (CVPR)* 10023-10033. doi: 10.1145/3581783.3611820.

Cormier, Kearsy, Fenlon, Jordan, Gulamani, Sannah, and Sandra Smith. 2015. "BSL Corpus Annotation Conventions, v. 2.0". *Deafness, Cognition and Language Research Centre*, University College London. http://www.bslcorpusproject.org/wp-content/uploads/BSLCorpus_AnnotationConventions_v2_-Feb2015.pdf (Accessed on Mar. 26, 2024).

Crasborn, Onno, Mesch, Johanna, Waters, Dafydd, Nonhebel, Annika, van der Kooji, Els, Woll, Bencie, and Brita Bergman. 2007. "Sharing sign language data online: Experiences from the ECHO project." *International Journal of Corpus Linguistics*, 12(4): 535-562. doi: 10.1075/ijcl.12.4.06cra.

Cuxac, Christian. 2000. *La langue des signes française (LSF): les voies de l'iconicité*. Ophrys, Paris, France.

da Rocha Costa, António Carlos, and Graçaliz Pereira Dimuro. 2001. "Supporting Deaf Sign Languages in Written Form on the Web". https://www.signwriting.org/forums/software/swml/swml01.html (Accessed on Dec. 4, 2023).

Doan, Patrick, Dominique Boutet, Adrien Contesse, Claudia S. Bianchini, Claire Danet, Morgane Rébulard, Jean-François Dauphin, Léa Chevrefils, Chloé Thomas, and Mathieu Reguer. 2019. "Handling Sign Language Handshapes Annotation with the Typannot Typefont". *CogniTextes* 19. doi: 10.4000/cognitextes.1401.

Duarte, Amanda, Shruti Palaskar, Lucas Ventura, Deepti Ghadiyaram, Kenneth DeHaan, Florian Metze, Jordi Torres, and Xavier Giro-i-Nieto. 2021. "How2Sign: A Large-scale Multimodal Dataset for Continuous American Sign Language". In *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR).* doi: 10.48550/arXiv.2008.08143.

Fabbretti, Daniela and Elena Pizzuto. 2000. "Dalle mani alla carta: aspetti teorici e metodologici della notazione della Lingua Italiana dei Segni". In *Rassegna di Psicologia* XVII(2): 101-122.

Fenlon, Jordan, Adam Schembri, and Kearsy Cormier. 2018. "Modification of Indicating Verbs in British Sign Language: A Corpus-Based Study". *Language 94*(1): 84-118. doi: 10.1353/lan.2018.0002.

Forster, Jens, Christoph Schmidt, Thomas Hoyoux, Oscar Koller, Uwe Zelle, Justus Piater, and Hermann Ney. 2012. "RWTH-PHOENIX-Weather: A Large Vocabulary Sign Language Recognition and Translation Corpus". In *Proceedings of the International Conference on Language Resources and Evaluation.* 3785-3789.

Ghosh, Abhigyan, and Mamidi Radhika. 2022. "English To Indian Sign Language: Rule-Based Translation System Along With Multi-Word Expressions and Synonym Substitution". In *Proceedings of the 19th International Conference on Natural Language Processing (ICON)* 123-127. New Delhi, India.

Gianfreda, Gabriele, Giulia Petitta, Claudia S. Bianchini, Alessio Di Renzo, Paolo Rossini, Tommaso Lucioli, Barbara Pennacchi, and Luca Lamano. 2009. *Dalla modalità faccia-a-faccia a una lingua scritta emergente: nuove prospettive su trascrizione e scrittura della Lingua dei Segni italiana (LIS).* Edited by C. Consani; C. Furiassi; F. Guazzelli. https://hal.science/hal-02366646/document (Accessed on Apr. 16, 2024).

Hanke, Thomas. 2004. "HamNoSys - Representing Sign Language Data in Language Resources and Language Processing Contexts". Workshop on the Representation and processing of Sign Languages. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, LREC 2004.

Johnston, Trevor. 2008a. "From archive to corpus: transcription and annotation in the creation of signed language corpora". In *Proceedings of Pacific Asia Conference on Language, Information and Computation* 106-131.

Johnston, Trevor. 2008b. "Corpus linguistics and signed languages: no lemmata, no corpus". Edited by O. Crasborn, E. Efthimiou, T. Hanke, E. Thoutenhoofd and I. Zwitserlood. In *Proceedings of the 5th Workshop on the representation and processing of signed languages: construction and exploitation of sign language corpora* 82-87. Paris, France.

Johnston, Trevor, and Adam Schembri. 2007. *Australian Sign Language (Auslan): An Introduction to Sign Language Linguistics*. Cambridge: Cambridge University Press.

Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Khan, Nabeel S., Adnan Abid, Kamran Abid, Uzma Farooq, Shoaib M. Farooq, and Jameel Hamza. 2021. "Speak Pakistan: Challenges in Developing Pakistan Sign Language using Information Technology". *Journal of South Asian Studies* 30 (2): 367-379.

Kooij, Els van der, Onno Crasborn, and Johan Ros. 2006. "Manual Prosodic Cues: PALM-UP and Pointing Signs". *TISLR 9*, Florianopolis, Brazil.

Kouremenos, Dimitris, Klimis Ntalianis, and Stefanos Kollias. 2018. “A novel rule based machine translation scheme from Greek to Greek sign language: Production of different types of large corpora and language models evaluation”. *Computer Speech & Language* (51): 110-135.

Liddell, Scott K. 2003. *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.

Lin, Chin-Yew. 2004. “ROUGE: A Package for Automatic Evaluation of Summaries”. In *Proceedings of the Workshop on Text Summarization Branches Out (WAS 2004)*. 74–81. Barcelona, Spain.

Lucioli Tommaso, Luca Lamano, and Gabriele Gianfreda. 2008. “I segni sulla carta: analisi e riflessioni sui primi testi di LIS scritta”. Edited by C. Bagnara, S. Corazza, S. Fontana & A. Zuccalà. *I Segni parlano. Prospettive di ricerca sulla Lingua dei Segni Italiana.* 148-158. Franco Angeli. Milano, Italy.

Marchetti, Carol, Susan Foster, Gary Long, and Michael Stinson. 2012. “Crossing the Communication Barrier: Facilitating Communication in Mixed Groups of Deaf and Hearing Students”. *Journal of Postsecondary Education and Disability* 25(1): 51–63. Rochester Institute of Technology.

McNeill, David. 1992. *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.

Montemurro, Kathryn, and Diane Brentari. 2018. “Emphatic Fingerspelling as Code-Mixing in American Sign Language”. In *Proceedings of the Linguistic Society of America* 3(1): 61:1–13.

Moryossef, Amit, and Yoav Goldberg. 2021. {S}ign {L}anguage {P}rocessing. https://sign-language-processing.github.io/ (Accessed on Mar. 18, 2024).

Moryossef, Amit, Mathias Müller, Anne Göhring, Zifan Jiang, Yoav Goldberg, and Sarah Ebling. 2023. “An Open-Source Gloss-Based Baseline for Spoken to Signed Language Translation”. In *Proceedings of the Second International Workshop on Automatic Translation for Signed and Spoken Languages*. 22-33. Tampere, Finland.

Mukushev, Medet, Arman Sabyrov, Madina Sultanova, Vadim Kimmelman, and Anara Sandygulova. 2022. “Towards Semi-automatic Sign Language Annotation Tool: SLAN-tool”. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*. 159-164. Marseille, France.

Muller, Mathias, Zifan Jiang, Amit Moryossef, Annette Rios, and Sarah Ebling. 2022. “Considerations for meaningful sign language machine translation based on glosses”. ArXiv preprint: arXiv:2211.15464.

Neves, Carolina, Luísa Coheur, and Hugo Nicolau. 2020. “HamNoSyS2SiGML: Translating HamNoSys Into SiGML”. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*. 6035-6039. Marseille, France.

Núñez-Marcos, Adrián, Olatz Perez-de-Viñaspre, and Gorka Labaka. 2023. “A survey on Sign Language machine translation”. *Expert Systems with Applications* 213(2): 118993. doi: 10.1016/j.eswa.2022.118993.

Ouargani, Younes, and Noussaima El Khattabi. 2023. “Advancing Text-to-GLOSS Neural Translation Using a Novel Hyper-parameter Optimization Technique”. ArXiv preprint: arXiv:2309.02162.

Padden, Carol A. 1998. “The ASL Lexicon”. *Sign Language & Linguistics* 1(1): 39-60.

Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. “BLEU: a method for automatic evaluation of machine translation”. *ACL-2002: 40th Annual meeting of the Association for Computational Linguistics* 311–318. doi: 10.3115/1073083.1073135.

Patrie, Carol J., and Robert E. Johnson. 2011. *Fingerspelled Word Recognition Through Rapid Serial Visual Presentation*. San Diego, CA: DawnSign Press.

Pennacchi Barbara, Elena Pizzuto, and Paolo Rossini. 2001. “La LIS: trascrizione e scrittura”. In *Seminari ISTC-CNR/ISSR*. Ciclo seminari 2000-2001, Roma, Italy.

Pfau, Roland. 2011. “A point well taken: on the typology and diachrony of pointing”. In *Deaf around the world: the impact of language*. Edited by G. Mathur and D.J. Napoli. 144-163. Oxford: Oxford University Press.

Pizzuto, Elena, Paolo Rossini, and Tommaso Russo. 2006. “Representing signed languages in written form: Questions that need to be posed”. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation* 1-6. AQ368 Genova, Italy.

Porta, Jordi, Fernando Lòpez-Colino, and Javier Tejedor. 2014. “A rule-based translation from written Spanish to Spanish Sign Language glosses”. *Computer Speech & Language* 28(3): 788-811. doi: 10.1016/j.csl.2013.10.003.

Prillwitz, Siegmund. 1989. “Hamnosys. version 2.0. Hamburg notation system for sign languages. An introductory guide”. *International studies on sign language and the communication of the deaf*. 5. Hamburg: Hamburg Signum Press.

Rastgoo, Razieh, Kiani Kourosh, Sergio Escalera, and Mohammad Sabokrou. 2021. “Sign Language Production: A Review”. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* 3446-3456. doi: 10.1109/CVPRW53098.2021.00384.

Rastgoo, Razieh, Kiani Kourosh, Sergio Escalera, Vassilis Athitsos, and Mohammad Sabokrou. 2022. “All You Need In Sign Language Production”. doi: 10.48550/arXiv.2201.01609.

Rathmann, Christian, and Mathur, Gaurav. 2011. “A Featural Approach to Verb Agreement in Signed Languages”. *Theoretical Linguistics* 37(3-4): 197-208. doi: 10.1515/thli.2011.015.

Sandler, Wendy. 2012. “The Phonological Organization of Sign Languages”. *Language and Linguistics Compass* 6 (3): 162-82. doi: 10.1002/lnc3.326.

Saunders, Ben, Necati C. Camgoz, and Richard Bowden. 2020. “Progressive Transformers for End-to-End Sign Language Production”. Edited by Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm. *Computer Vision - ECCV 2020, Lecture Notes in Computer Science* 687-705, Cham: Springer.

Schembri, Adam, Kearsy Cormier, and Jordan Fenlon. 2018. “Indicating Verbs as Typologically Unique Constructions: Reconsidering Verb 'Agreement' in Sign Languages”. *Glossa: A Journal of General Linguistics* 3 (1): 89. doi: 10.5334/gjgl.468.

Slobin, Dan, Nini Hoiting, Michelle Anthony, Yael Biederman, MarlonKuntze, Reyna Lindert, Jennie Pyers, Helen Thumann, and Amy Weinberg. 2001. “Sign language transcription at the level of meaning components: The Berkeley Transcription System (BTS)”. *Sign Language & Linguistics* 4(1-2): 63-104. doi: 10.1075/sll.4.12.07slo.

Slonimska, Anita, Asli Özyürek, and Olga Capirci. 2020. “The role of iconicity and simultaneity for efficient communication: The case of Italian Sign Language (LIS)”. *Cognition*, 200:104246. Epub Mar. 18, 2020. doi: 10.1016/j.cognition.2020.104246.

Stokoe, William C. 1960. “Sign language structure: an outline of the visual communication systems of the American deaf”. In *Studies in Linguistics, Occasional Papers*, 8. Buffalo, NY: Department of Anthropology and Linguistics, University of Buffalo.

Stokoe, William C., Dorothy Casterline, and Carl Croneberg. 1965. *A Dictionary of American Sign Language on Linguistic Principles*. Washington, DC: Gallaudet College.

Stoll, Stephanie, Necati C. Camgoz, Simon Hadfield, and Richard Bowden. 2018. “Sign language production using neural machine translation and generative adversarial networks”. In *Proceedings of the 29th British Machine Vision Conference (BMVC 2018)*. British Machine Vision Association. doi: 10.1007/s11263-019-01281-2.

Stoll, Stephanie, Necati C. Camgoz, Simon Hadfield, and Richard Bowden. 2020. “Text2sign: towards sign language production using neural machine translation and generative adversarial networks”. *International Journal of Computer Vision* 128: 891–908.

Supalla, Ted. 1983. "The Classifier System in American Sign Language. Noun Classes and Categorization". In *Proceedings of a symposium on categorization and noun classification*. 181-214. Eugene, Oregon. doi: 10.1075/tsl.7.13sup.

Sutton, Valerie, and Richard Gleaves. 1995. *SignWriter - The world's first sign language processor*. Center for Sutton Movement Writing. La Jolla. Ca.

Valeri, Vincenzo. 2003. *Per una scrittura della Lingua Italiana dei Segni*. Aracne. Roma, Italy.

Vauquois, Bernard. 1968. *A survey of formal grammars and algorithms for recognition and transformation in machine translation*. IFIP Congress 1968.

World Health Organization. 2021. World report on hearing. World Health Organization. https://iris.who.int/handle/10665/339913. License: CC BY-NC-SA 3.0 IGO.

Yin, Kayo, and Yesse Read. 2020. "Better sign language translation with STMC transformer". In *Proceedings of the 28th International Conference on Computational Linguistics* 5975-5989. Barcelona, Spain (Online). doi: 10.18653/v1/2020.coling-main.525.

Yin, Kayo, Amit Moryossef, Julie Hochgesang, Yoav Goldberg and Malihe Alikhani. 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* 1: 7347-7360. doi: 10.18653/v1/2021.acl-long.570.

Zhou, Hao, Wengang Zhou, Weizhen Qi, Junfu Pu, and Li Houqiang. 2021. "Improving Sign Language Translation with Monolingual Data by Sign Back-Translation". *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 1316-1325. doi: 10.1109/CVPR46437.2021.00137.

Zwitserlood, Inge, Margriet Verlinden, Johan Ros, and Sanny Van Der Schoot. 2005. "Synthetic signing for the deaf: Esign". In *Proceedings of the Conference and workshop on assistive technologies for vision and hearing impairment (CVHI)*. Granada, Spain.