

Italian Sign Language (LIS) and Natural Language Processing: an Overview

Sabina Fontana^[0000-0003-3083-1676] and Gaia Caligiore^[000-0002-7087-1819]

¹ University of Catania, Department of Humanities, Italy
sabina.fontana@unict.it

² University of Catania, Department of Humanities, Italy
gaia.caligiore@phd.unict.it

Abstract. The past decade has seen an increase in studies conducted on the interaction between NLP and sign languages. In this paper, we mainly focus on LIS while discussing the current state of the art, possible future developments and the ethical implications of this growing research context. In the history of NLP, human/computer interaction has been mainly based on the transcription of spoken languages. We investigate how existing resources for spoken language processing can be applied to SLs and combined with language-specific tools, providing examples of recent resources. We discuss novel strategies for sign transcription that consider both the need for standardized writing forms to enable NLP, as well as the language-specific features of SLs that are conveyed through the visual-manual channel. Deaf contributors are fundamental within this research. When NLP and SLs interact, we find a shift from a user-centric approach towards a user-based one to be essential: Deaf end-users of the resulting resources thus become part of the designing process¹.

Keywords: Italian Sign Language (LIS), Natural Language Processing, Translation.

1 Deafness and Deaf Communities

Different degrees of deafness affect millions of people around the world. The Italian ISTAT, National Institute for Statistics, states that in 2019, a total of 3 million people (5,2% of Italy's population) suffered from hearing loss [1]; showing the growth of hearing-impaired people who, in 1999, were estimated to be 877.000 [2]. Outside of Italy,

¹ The present paper has been co-authored by Sabina Fontana and Gaia Caligiore. Sabina Fontana developed section 1, subsections 2.1 and section 4. Gaia Caligiore developed subsections 2.2 to 2.6 and section 3. After proving feedback on the sections written by the co-author, the paper was jointly revised.

the WHO (World Health Organization) predicts that by 2050 nearly 2.5 billion people in the world will have some degree of hearing loss [3].

Whether they are born deaf or lose their hearing in their early infancy, due to illness or an accident, individuals with different levels of deafness (moderate, severe, profound) will be facing difficulties in communicating with hearing speakers. Such difficulties are of different kinds, depending on whether deafness is experienced in early infancy during language development or if it is acquired at an adult age, or if deaf children are exposed to a sign language in early infancy or not. About 10% of the deaf population is born deaf to deaf parents. The remaining 90% have hearing parents who do not know a sign language. This raises an issue concerning the transmission of sign languages, which tends to occur in a horizontal (among peers) rather than vertical way (from generation to generation). Recently, mainstream and technology have further influenced the educational path of deaf children and have led to delay or deny access to a sign language. Families tend to prefer normalization, choosing a spoken language and resisting bilingualism. Consequently, sign language is accessed and learnt very often at an adult age. Consequently, Deaf signers may acquire different levels of proficiency in spoken language and sign language [4].

Sign languages and spoken languages vastly differ since they employ different modalities: auditory-oral and visual-manual. For this reason, the bilingualism of Deaf signers is often referred to as bimodal [5]. Deaf people's bilingualism has some specific features. Firstly, it is bimodal [6] because they learn or acquire two languages exploiting different modalities through different paths: spoken language through speech therapy starting from early infancy and sign language through exposure, rarely during infancy, more frequently at a young age or even later in their lives. The two languages are used alternately but are in contact, so they influence each other while occurring simultaneously, since there are no serial order constraints [7]. Secondly, it is generally an unbalanced bilingualism: even if learnt later, the sign language appears to be their natural language and skills in spoken language are rarely comparable to natives². Lastly, the two languages have different sociolinguistic statuses. On the one hand, the spoken language – that is the majority language – is widely shared, institutional and used in education, in media communication and many other formal contexts. On the other hand, the sign language is a minority language that has been long stigmatized and only recently is being used in formal contexts, but still not enough in education. Although many schools and universities provide support in sign language through professional interpreters or communication assistants, there are only a few bilingual sign/spoken language schools in Italy (in Biella, Cossato and Rome) where LIS is studied like any other subject and taught to all students [8].

It is important to highlight that this condition of bilingualism influences the usage and the perception of sign languages from Deaf users [9]. It also influences the research work that must be carried out. It is fundamental to keep in mind that the target group is far from being homogeneous as it consists of people who have vastly different

² We are not using the categories ‘first language’, ‘second language’, and ‘mother tongue’ as they do not exactly mirror the specificities of deaf bilingualism.

experiences of Deafness and language acquisition. The experience of Deaf signers who present bimodal bilingualism will be different from that of individuals who lost their hearing later in life and are mainly faced with communication problems related to access rather than comprehension of a spoken language.

1.1 The Italian Sign Language (LIS)

As for spoken languages, when analyzing a sign language, adopting an interlinguistic perspective and a broader, more international frame of reference, will surely provide a rounder understanding of linguistic and social phenomena. Given our research interests, in this paper we will mainly focus on LIS. Therefore, all images and mentioned resources will be looked at as part of the LIS framework.

As for the signing population, the database *Ethnologue* [10] lists 121 sign languages used by 70 million Deaf people with 60.000 users of LIS. But what are the characteristics that determine the basic traits of LIS? Studying sign language from a phonocentric perspective can lead to the research of the same categories that define spoken languages. While this has been the preferred strategy in the past, it is fundamental to keep in mind that the different modalities (visual-manual/auditory-vocal) will call for different descriptions of a language that may not mirror each other. For this reason, we like to start describing LIS by observing how gestuality is systematically organized to create meaning. LIS employs the visual-manual modality, involving both manual and non-manual elements in the construction of signs. Manual elements are usually defined using the following major parameters: the shape one's hand (or hands) acquire while performing a sign, its place of articulation (also referred to as 'location'), the way hands move in space and hand orientation, i.e., the position of the palm of the dominant hand [11]. Non-manual elements play an equivalently crucial role in the construction of meaning and include head and body movements, facial expressions and mouth gestures [11]. These elements can simultaneously convey semantic information or have pronominal value during role-shift, a complex process that allows signers to 'become' the person, animal or object they are representing.



Fig. 1. The sign 'TU?' ('YOU?') in LIS. The man depicted in the figure is raising his eyebrows [12] and using a specific configuration of the lips. In doing so, he is creating the interrogative form of the sign. This example shows how non-manual elements allow us to distinguish between statements and questions.

The systematic creation of meaning in LIS can generate signs with different degrees of arbitrariness or iconicity. Gestuality is an essential aspect of human communication: if hands become the core elements of a language, there will be a continuity between gestuality and signs, which originates iconic phenomena. The presence of iconicity does not exclude arbitrariness. The spontaneous evolution of an arbitrary sign is arbitrary and unpredictable and takes place within a community that is impacted by specific cultural, social and geographical influences.

More than half a century has passed since the publication of the first methodological studies on sign languages. During this time, a shift in linguistics has slowly taken place, leading to a transformation in the perception of sign languages that are now universally accepted as natural ones and have been legally recognized as official languages in numerous countries. Changes in attitude and increased visibility forced signers to think upon their language and define a notion of correctness. At the same time, when sign languages became visible and accessible to a larger Deaf and hearing public (also through professional interpreting services in TV and other official contexts), Deaf people realized that it lacked many lexical items and functions to meet the different communicative needs [13].

The recognition of sign languages has been gradual but steady and has seen a fast development in the last 20 years. Within the Italian political and social context, the 2020 pandemic made LIS (*Lingua dei Segni Italiana*–Italian Sign Language) increasingly visible and public, as interpreters translated presidential speeches and conferences [14]. This newly acquired visibility–together with the widespread use of social media by Italian Deaf people–only sped up the ongoing process of standardization. At the same time, as LIS becomes more prominent and is used in more contexts, linguistic growth becomes necessary, leading to a rapid expansion of its vocabulary. Furthermore, the pandemic highlighted the hardship of Deaf people in contexts where face-to-face communication is limited, culminating in the official recognition of LIS and LIST (Tactile Italian Sign Language) by the Italian Government in May of 2021 [15].

As for the analysis of LIS structures, the main issue is that, up until the end of 2020, there were no grammars that extensively described its structural phenomena. This is due to different factors. LIS was first described in 1987 by Virginia Volterra [16]. At the time, the objective was the ‘dignification’ of the language, leading to a pursuit of the same parameters that define spoken languages within an assimilationist perspective. These first stages of LIS studies were characterized by a search for categories such as phonemes or minimal couples. After the first decade of research on LIS, signers themselves started using LIS in more formal environments [17]: the removal of LIS from the domestic and informal contexts led to an expansion and standardization of the language, which is still ongoing. With that came the newfound awareness of Italian signers who now see their language as a vehicle of Deaf pride and object in need of protection and preservation.

We take it for granted that sign languages are natural, meaning that they evolve spontaneously within a community. They are different from spoken languages in that they are not a manual translation of spoken languages and do not have standard written forms. Since the members of these communities are Deaf, they will make use of their bodies to

convey meaning. As we have said before, at present, sign languages are considered ‘oral’ in that there is no formal system for their transcription. As bimodal bilinguals, signers usually rely on the written form of spoken languages. This has largely influenced research on sign languages, leading to the use of glosses for sign transcription: the translation of the sign into spoken language written in all capitals. Of course, an external writing system that relies on translation does not convey the expressiveness of signs, hence the centrality of the need to move beyond glosses in sign language studies. Within this context, several attempts have been made at designing an independent signing system, the most successful being SignWriting [18].

2 Natural Language Processing and Italian Sign Language: Issues and Resources

We focus our attention on the possible contacts between NLP and sign languages to facilitate interactions between Deaf and hearing people and reflect on the practical and ethical challenges researchers are faced with when these two worlds come into contact. The sign language we focus on is Italian Sign Language (LIS). However, the observations we make in this paper refer to topics that lay at the core of all sign languages. We also make some observations on the implications of translating automatically sign languages into vocal languages and vice-versa.

In this section, we first focus on linguistic issues to be considered when working with LIS. We then move on to the current state of the art, describing projects for text to video synthesis in the Italian context. After that, we discuss available options for glossing LIS. We also provide Italian and international examples of different strategies employed to build datasets by converting text into sign language and sign language videos into text, through processes of manual and automatic recognition.

2.1 Linguistic and Ethical Issues

In shaping a dataset, different types of issues must be taken into consideration. First, we consider those issues that are related to the specific characteristics of LIS. We mentioned that sign languages should be looked at and analyzed using independent categories that do not necessarily mirror those of spoken languages. However, given that a computational analysis must go through a written message, it is necessary to reflect upon sign language transcription, its linguistic, social and ethical implications. Like other sign languages, LIS is an oral language and does not have a standardized writing system, given the visual-manual modality it employs and the fact that it relies on written Italian as an external writing system. An international, language-specific transcription system is SignWriting [18] which consists of a database of iconic symbols representing orientation and handshape of one or both hands, facial expression, movements, contact and location of the sign. These symbols are combined to represent signs. Due to its iconic nature, SignWriting allows for a detailed and simultaneous representation of the multilinearity of signed discourse but, despite the positive response from researchers, it has failed to

become the annotation system of choosing for signers. As discussed in subsection 1.1., only recently sign languages have been analyzed, described and recognized as natural languages. For these reasons, some signs could have not been developed yet. From a social standpoint, dataset collection should involve Deaf people not only as informants but as participants in that their living experience and linguistic knowledge is crucial for the development of the system that should account for their needs. The contribution of Deaf researchers is also important to define the setting for data elicitation to create a more natural dataset. Very often, when data are elicited in a very artificial setting, Deaf informants may shift to spoken language or mixed spoken and signed language varieties, i.e., contact signing [19]. We come to the second issue that should be considered in collecting datasets: the necessity to account for the variability deriving from different ages of acquisition and skills in spoken and sign languages. Ideally, to correctly identify linguistic structures, datasets should be collected in a naturalistic setting and involve Deaf people as participants.

2.2 Automatic Processing

Different issues arise once automatic processing is implemented on the language. Once data is obtained, researchers will be faced with the need for data annotation. A basic requirement is, of course, a readable transcription. This requirement is problematic for sign languages since current methodologies rely primarily on word labels: translations taken from spoken languages [20]. For this reason, in this section we also discuss tasks of word segmentation, reflecting on the issues one comes across when combining it with LIS. Word segmentation is a basic step, fundamental to sign annotation. When it comes to spoken languages, the content of each segment will naturally be one word, usually divided by the remaining string of written language by a space delimiter [21]. However, this step is not as straightforward when it comes to sign languages since it calls for the recognition of the beginning and end of a sign as well as the design of a transcription strategy. For this reason, gloss annotation is one of the main and most time-consuming issues [22].

2.3 State of the Art

Several models have been used to automatically transfer information from sign language to spoken language and vice versa. The general research problem is sign language recognition, as well as the processes leading to it [23]. Sign language recognition is concerned with the identification, segmentation and definition of signs. Given the multimodality and multilinearity of sign languages, data must be collected through video or images. To achieve said recognition, multi-channel approaches have been used, combining hardware-based or software-based resources [24]. In the past year, among the recent technologies employed for data collection for sign recognition, we find camera images [25] – in some instances acquired with RGB and RGB-D sensors [26] – fused with radar sensor data [27].

2.4 Text to Video Synthesis Projects for LIS

Within the Italian context, different projects have been carried out on the application of NLP and MT on LIS. We will now describe two related projects developed in the past years. The LIS4ALL project (2012–2014) [28]: funded in 2012 by Piedmont to create a prototype system for ITA-LIS automatic translation, providing a service for displaying information in LIS on mobile devices in train station terminals. The output was LIS utterances signed by an animated interpreter [20]. The LIS4ALL designers found that a small number of templates covered most announcements [29] and thus built three regular expressions that matched three templates. Following a process of sentence simplification, non-mandatory components (for example deictic signs) were not translated into LIS. The project resulted in the translation of 63 announcements. The main sources of error identified during translation were lexical gaps and the inability to handle the doubling of subject [29]. Given the novelty of the project, central aspects of LIS, such as non-manual elements and the morphological aspects of sign movement and location, were not considered in the project.

The LIS4ALL system was based on an existing one developed in the context of the ATLAS project for the translation of weather forecasts [30–31] through the creation of a lexicon of 2350 signs [30]. Moreover, the translation strategies used for LIS4ALL were interlingua rule-based and statistical translation [28], an evolution of the strategies that had been employed in the ATLAS project (Automatic Translation into sign LAnguageS, 2009–2012) for the description and identification of the relations between signs. For ATLAS, glossed LIS utterances had been analyzed using a set of lexical items and several combinatorial rules [28] that were elaborated by a LIS generator that ‘[...] builds a tree representing the generic LIS lexical items and some generic syntactic relations among them [...]’ [28]. Additionally, a set of values was associated with each sign gloss and provided information on the database name of the sign, its ID, the number of hands used to perform the sign and the part of speech of the sign [31] with the final aim of obtaining a signed target text performed by an avatar.

2.5 Going Beyond Glosses: Available Options for LIS

Traditionally, signs have been transcribed using glosses, i.e., a translation of the sign into the spoken target language. Therefore, in the context of sign transcription, if a LIS signer is describing something that happened involving something their dog did at home, we will surely have to transcribe ‘CANE’ (DOG) and ‘CASA’ (HOUSE). As can be inferred, despite their widespread use, glosses are flattening to the complexity of LIS. In fact, to a non-signer, the mentioned glosses provide no information on the production of the sign itself, only specifying its meaning. Additionally, signs may vary for many reasons, such as the geographical origin of a signer. Therefore, glosses make it virtually impossible for non-signers to obtain information on the sign once removed from the signed context.

Within the Italian framework, different strategies have been adopted to work around this issue. Generally, the most successful solution is the combination of video files and

univocal glosses. In the recently online-published resource *A Grammar of Italian Sign Language (LIS)* [32] the authors opted for a variation of gloss annotation by including videos of signers performing an isolated sign or an utterance in LIS, thus showing how signs combine and interact, as well as different variations of the same sign. The *Grammar* was published in the context of the SIGN-HUB project which aims at preserving and researching the linguistic, historical and cultural heritage of European Deaf signing communities with an integral resource [33].



- | | |
|---|---|
| a. GIANNI BUY HOUSE
'Gianni bought a house.' |  |
| b. GIANNI HOUSE BUY
'Gianni bought a house.' |  |

Fig. 2. An example of gloss writing taken from the *Grammar*. As can be seen, the authors included the gloss sequence as well as a translation into English. The translation into LIS of each utterance can be accessed by clicking on the hand-shaped icon to the right [32].

By associating glosses and video files, the *Grammar* certainly shows an aptitude towards a rework of the widespread glossing methodology. However, it does not collect its signs and utterances within a dictionary and, even if it did, the amount of data would be extremely limited. As for LIS dictionaries, at present, the only digital resources available to LIS researchers are the *Dizionario Bilingue Elementare della Lingua dei Segni Italiana (LIS)*³ [34] in its digital version, and the website SpreadTheSign [35–36].

The *Dictionary* is one of the most renowned and retrievable resources for LIS and it includes more than 2500 videos of signs performed by native signers. Each video is marked by a specific code that includes the translation of the sign into Italian and a sequence of numbers and/or letters. As mentioned, the richness of sign languages cannot be reproduced through capital letters [37] and might even lead to the creation of ambiguity, not providing information on the variation of the sign. For this reason, the *Dictionary* can be an excellent resource for LIS transcription since it provides an unambiguous code for each sign, thus creating an unequivocal association between the sign and its translation.

Despite its usefulness and the vastness of the resource, the *Dictionary* is limited in that it was concluded in 1992. For this reason, the multi-language dictionaries available on the website SpreadTheSign provide an invaluable contribution. The website is the result of an EU-funded project created for Deaf education and provides multilingual dictionaries for several sign languages from around the world.

³ Henceforth referred to as *Dictionary*.



Fig. 3. Image taken from the website SpreadTheSign [35]. The signer is performing the LIS sign HOUSE.

2.6 Existing Datasets for Sign Languages

In the previous section, we mentioned existing online dictionaries for LIS. We will now discuss different annotation tools and strategies used for other European and Northern American sign languages.

With regards to available resources for the segmentation and analysis of sign language videos, at present, ELAN is the most used resource in this field. It was initially released in 2000 by The Max Planck Institute for Psycholinguistics in Nijmegen, Netherlands, as a tool to annotate audiovisual files on different levels. ELAN is a useful tool for multimodality research on sign languages since it allows users to create time-aligned annotation levels where simultaneous information on manual and non-manual elements can be included [38].

Sign language translation requires glossing, which is a time-consuming, yet necessary, annotation process. Tokenization on the gloss level is widely used since a recognition of each sign that makes up an utterance facilitates the translation process. An example of a tokenized sign language corpus is the Swedish Sign language Corpus (SSLC), a resource developed at the Department of Linguistics of the University of Stockholm. The SSLC was compiled between 2009 and 2001 and included video-recorded conversations of 42 Swedish Sign Language Signers from the age of 20 to 82. The tokens collected for the SSLC are 33,600 [39]. The tokenization process was led on the ELAN platform. Annotators developed different levels (or tiers) to provide information on sign glosses taken from the Swedish Sign Language Dictionary [40], adding specifically developed tags and symbols to signal phenomena such as overlapping, merging, fingerspelling, or gesture-like sign. Another corpus partly annotated and tagged using ELAN is the British Sign Language Corpus Project (BSLCP) [41]. For the creation of the corpus, 249 BSL Deaf signers were recorded in conversational contexts. The ELAN software was used to provide information on what was being signed by the participants [42]. Within the context of the SIGN-HUB Project, Pfau [43] and other researchers from Germany, Italy, The Netherlands, Spain and Turkey, collaborated on a shared project for sign language annotation on ELAN. The goal was the inclusion of an annotation that went beyond translation or glossing, including aspects on non-manual productions and even comments from

annotators. As a result, more than 5 hours of footage were annotated by Deaf and hearing researchers.

All sign language data of the projects discussed in this section were annotated manually. As mentioned, annotation is a time-consuming process. To minimize the amount of time spent on it as well as facilitate the growth of large, annotated corpora for sign languages research, Drew and Ney [44] created a new interface for ELAN able to automatically recognize signs and annotate simultaneous tiers. The information included in the tiers is: glosses corresponding to the translations into spoken language as well as information on the features of each annotated sign. The richness of gloss annotation was designed to be modelled by users depending on their interests.

Given that most of the research on sign language recognition has been mainly focused on the tasks of gesture recognition, the linguistic qualities of sign languages have been overlooked. However, recent papers suggest a new approach to sign language translation problems. Sign language recognition uses contact [45] or vision-based systems [46]. The former is cumbersome, while the latter is mainly focused on identifying individual signs. Instead, real-world sign language recognition should be continuous to process the signing flow accurately. In 2018, Camgoz *et al.* [47] proposed the generation of spoken language translation from German sign language videos, mirroring steps of standard Neural Machine Translation which resulted in the PHOENIX4T dataset. In this dataset, gloss information makes up the data. Further development came with the introduction of sign language transformers [48] which can translate from spoken language sentences to a 3D skeleton [49].

To our knowledge, several attempts have been made at collecting sign language data by combining computer vision and information captured through gloves. Using the CopyCat system, Zafrulla *et al.* [50] collected 320 utterances from ASL signers. The tools for data collection used were a camcorder and colored gloves containing accelerometers providing information on acceleration, direction and rotation of hands. Another tool is the AcceleGlove, an electronic glove placed on a signer's hand and arm. The glove collected information on hand movement, orientation and location in relation to the body and recognized 176 signs in isolation [51] The AcceleGlove was combined with a gesture recognition toolkit [52] by McGuire *et al.* [45] to record 665 utterances in ASL to establish a pattern recognition framework to be expanded.

The idea of associating glosses taken from a set of values, together with additional information on the manual qualities of each sign, was developed in a 2020 master's thesis [53] and article [54]. The goal was the creation of a Universal Dependencies-compliant resource for the syntactic annotation of LIS, i.e., the first LIS treebank. To create this treebank, tasks of segmentation, annotation, POS tagging and parsing had to be carried out. Particular attention was paid to the inclusion of unambiguous information in the segmentation and annotation process. LIS videos were segmented on ELAN, after that, each sign was analyzed on different levels. The first tier provided an unambiguous gloss taken from the *Dictionary* or *SpreadTheSign*. The following tiers included information on sign location and Universal Dependencies POS tag. Utterances were then transferred into CoNLL-U format for the construction of dependency trees. During that step, gloss, sign location and POS tag were transferred in the columns, together with a translation

into Italian. The treebank can be found on GitHub [55]. The main issue encountered in this project is, once again, the lack of information on the annotation of non-manual elements. Annotating on ELAN, which allows for simultaneous viewing of video and annotation, is convenient. However, once we move to the level of syntactic annotation in CoNLL-U format, the video material is no longer visible. Therefore, all information that could be not codified in the tiers through word labels is virtually inaccessible.

As regards available datasets, a list of sign languages recognition datasets with information on data type and annotation systems, as well as related papers can be found in the bibliography [56].

3 Additional Challenges for LIS–IT and IT–LIS Translation

3.1 Subdivision of Manual and Non-Manual Elements

Sign transcription for data collection is not the only relevant problem to be faced for sign language processing. On the one hand, the manual elements that make up a sign in LIS have been divided into different parameters: handshape, place of articulation, orientation and movement. On the other hand, non-manual elements are equally important in utterance construction and sign disambiguation and include facial expression (movements of eyebrows, eye, mouth, nose), body posture and movements.

Once a methodology for sign transcription is identified, the second aspect to be considered is the identification of the relevant elements in sign construction. Can sign annotation be limited to only manual ones? Or are non-manual elements fundamental and cannot be taken out of the discourse? This issue holds a central spot in Italian and international contexts. Two main perspectives provide different answers to this issue. ‘Assimilationists’ show a tendency to focus on those aspects that draw sign languages closer to spoken ones, thus giving priority to ‘standard signs’, which are easier to define and mainly constituted by manual elements. ‘Non-assimilationists’ want to highlight the language-specific properties of sign languages, such as the relevance of non-manual elements [57].

Whichever approach one may choose, the search for relevance remains central. Ideally, manual and non-manual elements should be taken into consideration, thus providing information on every aspect of the sign. However – given the current technologies employed in this field and discussed in section 2.3., such as RGB cameras or radar sensors – manual elements hold a central position in the investigation, temporarily winning over non-manual ones.

3.2 Challenges of LIS Utterance Structure Description

Despite our non-assimilationist inclinations, it is undeniable that if the aim of a process of data collection and annotation is MT, a dismissal of the pre-existing POS tagging systems is counterproductive. For this reason, in this subsection, we examine recent assimilationist observations on LIS utterances, by taking into consideration the manual and non-manual levels. Utterances are based on signs, the entire body and non-manual

features such as facial expression, mouth actions, movements of the torso and eye gaze, which is used to point at, describe or depict the referent.

Sign languages have specific constructions: the structure of an utterance in LIS will not mirror that of spoken Italian. It has been observed that the unmarked order of signs in LIS is Subject-Object-Verb, describing LIS as a head-final language, where the most meaningful element is found in final position [58]. Utterances are in most cases much more complex as they function following pragmatic constraints based on iconicity and the multifaceted structures of LIS have not been comprehensively defined up to this point.

Non-manual elements represent the hardest challenge in the representation and recognition of sign languages for their non-segmentable nature. Facial expressions that also includes eye gaze and mouth actions convey relevant information in co-occurrence with signing that is hard to process as relevant but is crucial for the understanding of the meaning of the single sign and the utterance.

4 Conclusion

This paper explores some of the main challenges in the field of sign language processing. We have discussed the socio-political situation of the Deaf community and considered how, behind the label of Deafness, there can be different users with different needs. This highlights that the involvement of users is essential in every phase of research and development. When creating sign language datasets, Deaf people should be involved in collecting reliable data that represent sign language usage, making possible the creation of appropriate computational models, interface design and, finally, of the overall systems. Furthermore, the creation of datasets based on the processing of natural conversation and long utterances will be necessary to go beyond the state of the art. Another crucial step is annotation. The lack of a standard written form and the necessity of fluent signers who annotate to produce the machine-readable inputs for training algorithms, represent the main difficulties in applying NLP methods to sign languages. Only through a multidimensional approach that combines NLP with computer vision and radar-based technologies, as well as with the involvement of Deaf participants, could it be possible to design effective technologies that could have an impact both at the social and the linguistic level. On a social level, an effective translation could support the interactions between hearing and Deaf people when the interpreting service is not available, for example. On a linguistic level, sign language processing would play an important role in sign language acquisition and learning.

The proper development of this new technology will innovate along two main directions: technological and social. Under the technological aspect, the collaboration of different research approaches (radar technologies, artificial intelligence, and computer vision research) can open new insights in understanding the shaping of machines to respond to human beings' needs. On the other hand, it will have high benefits in terms of the inclusion of Deaf people.

References

1. Istituto Nazionale di Statistica: Conoscere il Mondo della Disabilità: Persone, Relazioni e Istituzioni. Istituto Nazionale di Statistica, Roma (2019).
2. Deafness and Hearing Loss <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>, last accessed 2021/10/02.
3. Malerba, D.: Sordità Percezione e Realtà nell'approccio pedagogico. Sapienza Università Editrice, Roma (2020).
4. Fontana, S., Mignosi, E.: Segnare, Parlare, Intendersi: Modalità e forme. Mimesis, Milano and Udine (2012).
5. Onofrio, D., Rinaldi, P., Caselli, M.C., Volterra, V.: Il Bilinguismo Bimodale dei Bambini Sordi: Aspetti Teorici ed Esperienze di Ricerca. Rivista di psicolinguistica applicata XIV (1), 25–42. (2014).
6. Pinto M. A., Volterra V.: *Bilinguismo lingue dei segni / lingue vocali: Aspetti educativi e psicolinguistici* – Sign languages, spoken languages and bilingualism: educational and psycholinguistics issues. Italian Journal of Applied Psycholinguistics VIII (2008).
7. Capek, C. M. *et al.*: Hand and Mouth: Cortical Correlates of Lexical Processing in British Sign Language and Speechreading English. *Journal of Cognitive Neuroscience* 20 (7), pp. 1220–1233 (2008).
8. Maragna S., Roccaforte M., Tomasuolo E.: *Una Didattica innovativa per l'apprendente sordo*. FrancoAngeli, Milano (2013).
9. Fontana S: Metalinguistic Awareness in sign language: epistemological considerations, in Pinto M.A., Rinaldi P. (eds) Metalinguistic Awareness and Bimodal Bilingualism: Studies on Deaf and Hearing Subjects, Italian Journal of Applied Psycholinguistics XVI(2) (2016).
10. Ethnologue Homepage, <https://www.ethnologue.com/>, last accessed 2021/10/08.
11. Scelzi, R.: Le component non manuali della LIS. Studi di Glottodidattica (1), 261–291 (2010).
12. Romeo, O.: Dizionario dei Segni. Zanichelli Editore S. p. A., Bologna (1991).
13. Fontana S., Corazza S., Boyes Braem P., Volterra V.: Language research and language community change: Italian Sign Language 1981–2013, in International Journal of the Sociology of Language, pp. 1–30. De Gruyter Mouton, Berlin (2015).
14. Tomasuolo E., Gulli T., Volterra V., Fontana S.: The Italian Deaf Community at the time of Coronavirus. *Frontiers in Sociology* (2021). DOI: <https://doi.org/10.3389/fsoc.2020.612559>
15. Italian Official Gazette, Decree Law of March 22nd 2021, n.41, https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2021-05-21&atto.codiceRedazionale=21A03181&elenco30giorni=false, last accessed 2021/10/01.

16. Volterra, V.: *La Lingua Italiana dei Segni. La comunicazione visivo-gestuale dei sordi*. Il Mulino, Bologna (1987).
17. Fontana, S., Roccaforte, M.: *Oltre l'approccio assimilazionista nella descrizione LIS. Quando la prassi comunicativa diventa norma*. In: *LII Congresso Internazionale di Studi della Società di Linguistica Italiana*, pp. 273–286. Bern (2019).
18. Di Renzo, A., Lamano, L., Luciola, T., Pennacchi, B., Ponzio, L.: *Italian Sign Language (LIS): can we write it and transcribe it with SignWriting?* In: Editor, F., Editor, S. (eds.) *2nd Workshop on the Representation and processing of Sign Languages: Lexicographic matters and didactic scenarios*, International Conference on Language Resources and Evaluation –LREC 2006, Genoa (2006).
19. Volterra V., Roccaforte M., Di Renzo A., Fontana S.: *Descrivere Lingue dei Segni. Una prospettiva sociosemiotica*. Il Mulino, Bologna (2019).
20. Antinoro Pizzuto, E., Chiari, I., Rossini, P.: *The Representation Issue and its Multifaceted Aspects in Constructing Sign Language Corpora: Questions, Answers, Further Problems*. In: *Proceedings of the LREC2008 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, pp. 150–150. ELRA, Marrakech (2009).
21. Mitkov, R.: *The Oxford Handbook of Computational Linguistics*. Oxford University Press, New York (2003).
22. Camgöz, N.C., Koller, O., Hadfield, S., Bowden, R.: *Multi-channel Transformers for Multi-articulatory Sign Language Translation*. (2020).
23. Manaris, B.: *Natural Language Processing: A Human-Computer Interaction Perspective*. *Advances in Computers* 47, 1–66 (1998).
24. Farooq, U., Rahim, M.S.M., Sabir, N., Hussain, A., Abid, A.: *Advances in machine translation for sign language: approaches, limitations, and challenges*. *Neural Computing and Applications* (2021).
25. Nobis, F., Geisslinger, M., Weber, M., Johannes, B., Lienkamp, M.: *A Deep Learning-based Radar and Camera Sensor Fusion Architecture for Object Detection*. (2020).
26. Huang, J., Wen-gang, Z., Qilin, Z., Houqiang, L., Weiping, L.: *Video-based Sign Language Recognition without Temporal Segmentation*. *AAAI* (2018).
27. Santhalingam, P., Du, R., Wilkerson, R., Al Amin, H., Ding, Z., Parth, P., Rangwala, H., Kushalnagar, R.: *Expressive ASL Recognition using Millimeter-wave Wireless Signals*. *2020 17th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 1–9. (2020).
28. Lombardo, V., Battaglino, C., Damiano, R., Nunnari, F.: *An Avatar-based Interface for the Italian Sign Language*. *International Conference on Complex, Intelligent and Software Intensive Systems, CISIS 2011*. Korean Bible University, Seoul (2011).
29. Battaglino, C., Geraci, C., Lombardo, V., Mazzei, A.: *Prototyping and Preliminary Evaluation of Sign Language Translation System in the Railway Domain*. *International Conference on Universal Access in Human-Computer Interaction*, pp. 229–350. (2015).

30. Mazzei, A.: Translating Italian to LIS in the Rail Stations. Proceedings of the 15th European Workshop on Natural Language Generation, ENLG. Association for Computational Linguistics, Brighton (2015).
31. Mazzei, A., Lesmo, L., Battaglino, C. Vendrame, M., Bucciarelli, M.: Deep Natural Language Processing for Italian Sign Language Translation. In: Baldoni M., Baroglio C., Boella G., Micalizio R. (eds) Advances in Artificial Intelligence, AI*IA 2013, vol 8249. Springer, Cham.
32. Branchini, C., Mantovan, L., A Grammar of Italian Sign Language (LIS). Published with open access at <https://www.sign-hub.eu/grammardetail/UUID-GRMM-e0adecd1-c01e-47ef-b2c0-c2d6a4ce45dc> (2020).
33. The SIGN-HUB Project Homepage, <https://www.sign-hub.eu/project>, last accessed 2021/10/04.
34. Radutzky, E.: *Dizionario bilingue elementare della lingua italiana dei segni. Oltre 2500 significati. Con DVD-ROM*. Kappa (1992).
35. SpreadTheSign Homepage, <https://www.spreadthesign.com/it.it/search/>, last accessed 2021/10/04.
36. Hilzensauer, M., Krammer, K.: A multilingual dictionary for sign languages: “SpreadTheSign”. The 8th annual International Conference of Education, Research and Innovation, ICERI2015, Seville (2015).
37. Chesi, C., Geraci, C.: Segni al computer. Manuale di documentazione della Lingua Italiana dei Segni e alcune applicazioni computazionali. Siena (2009).
38. Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., and Sloetjes, H.: ELAN: a professional framework for multimodality research. Proceedings of the Fifth International Conference on Language Resources and Evaluation, LREC’06. European Language Resources Association (ELRA), Genoa (2006).
39. Mesch, J., Wallin, L.: Gloss annotations in the Swedish Sign Language Corpus. *International Journal of Corpus Linguistics* (20), pp. 102–120 (2015).
40. *Svenskt teckenspråkslexikon* (Swedish Sign Language Dictionary), <http://teckensprakslexikon.su.se/>, last accessed 2021/10/08.
41. BSLCP Homepage, <https://bslcorpusproject.org/>, last accessed 2021/10/08.
42. Schembri, A.: British Sign Language Corpus Project: Open Access Archives and the Observer’s Paradox. (2015).
43. Pfau, R.: Annotation in ELAN, Version 1.1. SIGN-HUB Project, European Commission (2018).
44. Drew, P., Ney, H.: Towards Automatic Sign Language Annotation for the ELAN Tool. (2008).
45. McGuire *et al.*: Towards a One-way American sign language translator. In Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 620– 625. Seoul (2004).
46. Santhalingam, P.S., *et al.*: Expressive ASL Recognition using Millimeter-wave Wireless Signals. In 17th Annual IEEE International Conference on Sensing, Communication and Networking, pp. 1–9. SECON, Como (2020).

47. Camgoz, N.C., Hadfield, S., Koller, O., Ney, H., Bowden, R.: Neural Sign Language Translation. In: Conference on Computer Vision and Pattern Recognition, pp. 7784–7793. (2018).
48. Camgoz, N.C., Hadfield, S., Koller, O., Bowden, R.: Sign Language Transformers: Joint End-to-End Sign Language Recognition and Translation. In: Conference on Computer Vision and Pattern Recognition. (2020).
49. Saunders, B., Camgoz, N.C., Bowden, R.: Progressive Transformers for End-to-End Sign Language Production. In: European Conference on Computer Vision (ECCV) (2020).
50. Zafrulla *et al.*: American Sign Language Recognition with the Kinect. In: Proceedings of the 13th International Conference on Multimodal Interfaces, ICMI 2011. Alicante (2011).
51. Hernandez-Rebollar, J.L., Kyriakopoulos, N., Lindeman, R.W.: A New Instrumented Approach for Translating American Sign Language into Sound And Tex. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition. Seoul (2004).
52. Westeyn, T.L., Brashear, H., Atrash, A., Stamer, T.: Georgia tech gesture toolkit: supporting experiments in gesture recognition. In Proceedings of the 5th International Conference on Multimodal Interfaces, ICMI 2003. Vancouver (2003).
53. Caligore, G.: Universal Dependencies for Italian Sign Language: a treebank from the storytelling domain. University of Turin, Turin (2020).
54. Caligore, G., Bosco, C., Mazzei, A.: Building a treebank in Universal Dependencies for Italian Sign Language. Proceedings of Seventh Italian Conference on Computational Linguistics, CLIC-2020, vol. 2769, (2021).
55. Caligore, G., Bosco, C., Mazzei, A.: LIS-UD. <https://github.com/alexmazzei/LIS-UD.git>, (2020), last accessed 2020/10/06.
56. Sign language datasets http://facundoq.github.io/guides/sign_language_datasets/slr, last accessed 2021/10/03.
57. Scelzi, R.: Le Componenti non Manuali (CNM) della LIS. In: Studi di Glottodidattica, pp. 261–291. (2010).
58. Geraci, C.: L’ordine dei segni nella LIS (Lingua dei Segni Italiana). University of Milan, Milan (2002).