

Annotating a Fable in Italian Sign Language (LIS)

Serena Trolvi, Rodolfo Delmonte

Ca' Foscari University of Venice

Ca' Bembo - Department of Language Sciences

841996@stud.unive.it, delmont@unive.it

Abstract

This paper introduces work carried out for the automatic generation of a written text in Italian starting from glosses of a fable in Italian Sign Language (LIS). The paper gives a brief overview of sign languages (SLs) and some peculiarities of SL fables such as the use of space, the strategy of Role Shift and classifiers. It also presents the annotation of the fable “The Tortoise and the Hare” - signed in LIS and made available by Alba Cooperativa Sociale -, which was annotated manually by first author for her master’s thesis. The annotation was the starting point of a generation process that allowed us to automatically generate a text in Italian starting from LIS glosses. LIS sentences have been transcribed with Italian words into tables on simultaneous layers, each of which contains specific linguistic or non-linguistic pieces of information. In addition, the present work discusses problems encountered in the annotation and generation process.

Keywords: Italian Sign Language, annotation, fables, Natural Language Generation

1. Introduction

This paper presents work carried out for the automatic generation of a written text in Italian starting from glosses of a fable in Italian Sign Language (LIS).

Sign Languages (SLs) are languages used worldwide by Deaf¹ communities. They are natural languages, like spoken languages. However there is a crucial difference between them: communication occurs through the vocal-auditory channel in spoken languages; sign languages operate within a different modality instead and employ the visual-manual channel. This allows signers to use the whole upper body to communicate. In fact, SLs are not expressed using only the hands: non-manual articulators like head, eye gaze and shoulders play a crucial role too. Hence, it can be easily inferred that information is often conveyed simultaneously.

Signed narratives are very common in Deaf communities (Valli and Lucas, 2000) and fables are also popular. Fables are short or medium-length stories with a moral and they generally have specific characteristics in SLs that are usually not to be found in spoken languages like Italian.

In this paper we present the annotation method used for the fable “The Tortoise and the Hare” - signed in LIS and made available by Alba Cooperativa Sociale -, which was annotated manually by first author for her master’s thesis. It is about a tortoise and a hare that decide to have a race. As described in another paper (Delmonte and Trolvi, 2020), the annotation of the story was the starting point of a generation experiment that allowed us to automatically generate a text in Italian starting from LIS glosses. The story was annotated transcribing LIS information with Italian words into tables on simultaneous layers, each of which contains specific linguistic or non-linguistic pieces of information. This work also focuses on describing problems encountered during the annotation and generation process.

¹Capitalized letter D in the term “Deaf” refers to those deaf people who identify themselves as members of a signing community. The word “deaf” with lower case indicates only the condition of hearing loss (Woodward, 1972).

2. Sign Languages

2.1. General Overview

In contrast to what many people think, Sign Languages are full-fledged natural languages with their own phonology, morphology and syntax and not mere gestural systems without grammar. They are not dependent from spoken languages and are not universal: each country generally has its own sign language, and some countries have more than one (Caselli et al., 2006).

The same applies to Italian Sign Language (LIS), which is the sign language used by the Italian Deaf community, although it lacks official recognition at national level to date². Sign Languages also have specific characteristics and can simultaneously convey information by means of manual and non-manual components.

Manual components are signs realized using hands and can be divided into four phonological parameters:

- (i) handshape, which is the hand configuration;
- (ii) location, that is the location on the signer’s body or in the signing space, i.e. the area in front of the signer in which signs are realized;
- (iii) movement, which describes the hand movement during the realization of the sign;
- (iv) palm orientation, that refers to the direction in which the hand is turned to produce the sign (Stokoe, 1960; Battison, 1978).

²For linguistic research on LIS see Volterra et al. (1984), Volterra et al. (1986), Pizzuto et al. (1990), Corazza (1990), Pizzuto et al. (1995), Pizzuto and Corazza (1996), Geraci (2002), Volterra (2004), Zucchi (2004), Geraci (2006), Caselli et al. (2006), Branchini (2007), Geraci et al. (2008), Mazzoni (2008), Bertone (2009), Mazzoni (2009), Bertone and Cardinaletti (2009), Branchini and Donati (2009), Cecchetto et al. (2009), Zucchi (2009), Bertone (2011), Branchini and Geraci (2011), Mazzoni (2012), Branchini (2014), Donati et al. (2017).

In LIS 38 handshapes, 15 locations, 32 movements and 6 palm orientations have been identified (Volterra, 2004).

Non-manual components (or Non-Manual Markers, NMMs) include facial expressions and movements of head, eyebrows, eyes, cheeks, mouth, shoulders and postures. Some researchers consider NMMs to be the fifth phonological parameter (Valli and Lucas, 2000). They convey information together with manual signs and may have linguistic and non-linguistic functions (Corina et al., 1999; McCullough et al., 2005; Pfau and Quer, 2010; Herrmann, 2013).

Non-linguistic NMMs have an affective function and are used to convey emotions, whereas linguistic NMMs play specific grammatical roles. They are both realized through the same articulators, but they differ in scope, timing and involved muscles (Baker and Padden, 1978; Hickok et al., 1996; Corina et al., 1999; McCullough et al., 2005). In fact, linguistic facial expressions have a clear onset and offset, are timed to co-occur with specific constituent structures and involve restricted facial muscles (Baker-Shenk, 1983). On the contrary, affective expressions are gradual and vague, they have a more inconsistent onset and offset and are not accurately coordinated with signs. Furthermore, they involve a global activation of facial muscles (Liddell, 1978; Liddell, 1980; Emmorey, 1999; Wilbur, 2000; Herrmann, 2013).

Linguistic NMMs can be distinguished in lexical, adverbial and syntactic.

Lexical NMMs necessarily co-occur with specific manual signs. Adverbial NMMs modify adjectival and adverbial information. Syntactic NMMs determine the sentence type, mark topicalized constituents and are also realized in subordinate clauses (Pfau and Quer, 2010).

They can spread all over the sentence or co-occur with particular phrases. As for LIS, NMMs distinguish yes/no and WH questions, negative and imperative sentences, topic and focus phrases. Furthermore, they are realized in conditional and relative constructions.³

However, NMMs do not always occur in simultaneous conjunction with manual signs. Dively (2001) investigated in American Sign Language (ASL) the so-called "nonhanded signs" (NHSs), namely signs realized without using the hands. The author identifies eight NHSs (i.e. NHSs for "yes" and "no") as lexical items, on the basis of their form, meaning and function. She considers them free morphemes and points out that they are not universal.

Similar nonhanded signs have also been found in Polish Sign Language (Tomaszewski and Farris, 2010). However, Herrmann (2013) observes that NHSs seem to be used as discourse structural components and not as lexical elements.

³For further reading on these topics please see Franchi (2004); Geraci (2006); Branchini (2007); Branchini and Donati (2009); Cecchetto et al. (2009); Donati et al. (2017).

2.2. Some Characteristics of Fables in Sign Language

Fables are short or medium-length stories with a moral and are usually told to children. However, this does not mean that they are simple texts. On the contrary, they have a very complex structure.

As far as Sign Languages are concerned, signed narratives are very common in Deaf communities and fables are popular too. SL narratives - including fables - usually have at least three peculiarities, which are not used in the same way in spoken languages and are linked respectively to the use of space, Role Shift and classifiers.

2.2.1. Use of Space

Unlike spoken languages, which use the vocal modality, sign languages convey meaning simultaneously through the visual channel. Many signs are produced in the area in front of and at the side of the signer, which is the neutral signing space and extends from the top of the head to the waist and from shoulder to shoulder. According to Bertone (2009), the realization of the sign in a particular and definite point in the neutral space can have specific purposes. This is mostly due to two reasons. The former concerns agreement necessities: points in the space can be defined as morphemes that establish agreement between nouns and verbs. The latter is related to specification marking: if a point in the space is specific and definite, it is identified as a point that differs from other points in that space. Hence, points in the space may be used to locate referents or entities and create anaphoric references. Taking this into account, we can distinguish a referential, defined space from a neutral, undefined space⁴.

Let us now focus on the representation of the different places of articulation, which remains a problematic and widely discussed topic. In different contexts, points in the space do not usually preserve the same reference and do not usually convey the same meaning. In fact, they are defined in a particular situation by the signer, who always addresses an interlocutor. These points cannot be interpreted as fixed elements in the space, since they change depending on where signer and interlocutor are located, as noticed by Bertone (2009).

Figure 1 (Mac Laughlin, 1997) shows how the representation works. On this basis, we can outline some space features.

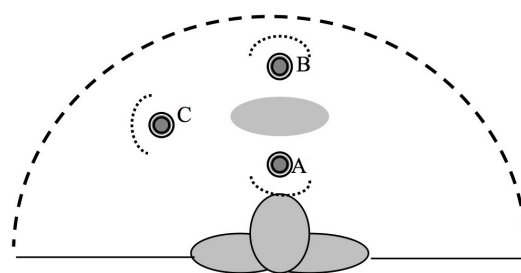


Figure 1: Top view of the signing space

⁴See Klima and Bellugi (1979); Padden (1990); Bahan (1996); Meier (1990); Liddell (1995); Liddell (2002) for further research.

Points in the space have two semantic features: proximal [+/-prox] and distal [+/-dist]. The former involves signer and interlocutor; the latter represents a third person or entity. Point A represents the nearest point to the signer [+prox]. Point B is farther from the signer and located between him and the interlocutor [-prox]. Point C refers to someone or something that is far from both signer and interlocutor [+dist]. As can be inferred, each point is specific - because it is bound to a referent - and definite - because it can be identified -. Consequently, an unspecified point in the space - e.g. "somewhere" - is undefined: [-dist] (Bertone, 2009).

2.2.2. Role Shift

Role Shift (RS) is a particular narrative strategy by which the signer adopts the perspective of another referent. It is a grammaticalized phenomenon widely used in many Deaf communities⁵, but not very common in spoken languages. It is characterized by specific body markers, which according to Mazzoni (2012) are:

- (i) precise position of the referents within the space, which becomes a real narrative setting;
- (ii) temporary interruption of eye contact with the interlocutor of the signer and change of direction of eye gaze towards the point of the space that is associated with the interlocutor of the embodied referent;
- (iii) body shift towards the point of the space that is associated with the embodied referent;
- (iv) change in head position;
- (v) facial expression associated with the embodied referent.

In addition, RS involves the displacement of indexical elements: first and second person pronouns do not refer to the signer and the interlocutor of the main context of utterance, but to those of the reported one. The interpretation of temporal and locative indexical is based on the derived context too (Quer, 2016; Schlenker, 2017a; Schlenker, 2017b).

Role Shift can be used to report a speech or thought of a referent or to reproduce his or her actions, so it can be divided into two varieties (Schlenker, 2017a; Schlenker, 2017b). However, the terminology for both phenomena is not consistent throughout the literature⁶. In this work, we adopt

⁵RS has been studied in ASL by Bahan and Petitto (1980), Padden (1986), Meier (1990), Lillo-Martin (1995), Lee et al. (1997); in Swedish Sign Language by Ahlgren (1990); in Danish Sign Language by Engberg-Pedersen (1995); in South African Sign Language by Aarons and Morgan (2003), in French Sign Language by Cuxac (2000) and many more. As for LIS, the reader is referred to Ajello (1997), Zucchi (2004), Mazzoni (2008), Mazzoni (2009), Mazzoni (2012).

⁶For instance, Metzger (1995) and Cormier et al. (2011) name RS "constructed action". Pfau and Quer (2010) and Lillo-Martin (2012) oppose "quotational" with "not-quotational uses of role shift". Herrmann and Steinbach (2012) call these categories "role shift" and "constructed action". Schlenker (2017a) and Schlenker (2017b) distinguishes between "attitude role shift" and "action role shift".

the terminology used by Herrmann and Pendzich (2018), namely "quotation role shift" (QRS) and "action role shift" (ARS).

Hence, QRS is the type of RS by which the signer reports words or thoughts of other referents.

ARS allows the signer to iconically reproduce actions, mannerisms and emotional states. It involves the use of the upper parts of the body (e.g. torso, head, eye gaze), including facial expressions and non-linguistic gestures.

2.2.3. Classifiers

Classifiers (CLs) are grammatical units used in many spoken languages to classify nouns according to specific semantic classes⁷. A CL is usually considered as an overt morpheme and indicates a peculiar feature of a referent, such as its shape or semantic category.

According to Mazzoni (2012), CLs are found in all studied sign languages⁸ and do not critically differ in their functions from classifiers in spoken languages. Classifiers in SLs consist of specific handshapes realized in combination with specific predicative roots - i.e. movements - and can convey different pieces of information simultaneously. As for LIS, the author suggests a model based on studies of Engberg-Pedersen (1993) on Danish Sign Language and Benedicto and Brentari (2004) on ASL.

Classifiers are classified into four semantic categories:

- (i) Whole Entity Classifiers, that include handshapes that represent the nominal referent as a whole;
- (ii) Handling/Instrument Classifiers, which indicate how the hand holds an object while holding and manipulating it;
- (iii) Extension and Surface Classifiers, that refer to a specific characteristic or feature of the nominal referent;
- (iv) Limb/Body Part Classifiers, which denote specific parts of a body, such as head, feet, eyes, but also paws, tail and horns.

These are combined with the following predicative roots:

- (i) Action or Movement Roots, which are morphemes describing the movement of the nominal referent, such as a person that moves from left to right;
- (ii) Manner or Imitation Roots, that represent an action or the type of movement of the referent;
- (iii) Position or Contact Roots, in which the combination movement-hold produced by the hand situates the referent in the space;
- (iv) Extension or Stative-Descriptive Roots, in which the hand movement shows the state of the referent or illustrates how referents are positioned in the space.

⁷For further reading on classifier systems in spoken languages, the reader is referred to Allan (1977), Craig (1992) and Aikhenvald (2000) among others.

⁸Please see Schembri (2003):28 – endnote 1 and Mazzoni (2012):35 for further research.

An example of classifiers is given in Figure 2 - which is taken from our fable -, in which the hands represent two Whole Entity Classifiers combined with Movement Roots.⁹ In that context, these classifiers represent the tortoise and the hare moving closer to each other and to an imaginary starting line, in order to start the race.



Figure 2: "Example of two classifiers in LIS"

From the above we can easily infer that classifiers are not standard, but rather variable structures. This also accounts for the fact that such signs are usually not reported in sign language dictionaries.

3. The Generation Experiment in a Nutshell

As already said, our experiment consists in the automatic generation of a text in Italian starting from glosses in LIS. Generation is carried out by our GENLIS generator, which is written in the programming language Prolog. As described in detail in another paper (Delmonte and Trolvi, 2020), the generation system is based on a set of algorithms that allow the generator to convert the glosses into text.

The signed text we chose is a fable - "The Tortoise and the Hare" - because it is a semantically and pragmatically difficult text to generate. The annotated text is organized into Discourse Units (DUs), each of which may contain more than one sentence and is associated with a unique turn and a unique speaker.

Annotation is the starting point of the process. Glosses are organized on different simultaneous layers. Glossed terms are then transformed into linear Prolog terms, which are one-layer strings composed of slots - each of which corresponding to a specific layer of glosses - to be fed as an input to the generator. In order to obtain a well-formed output text, input terms - i.e. information pieces in every DU - have to be tagged and processed separately, on the basis of their grammatical function and semantic features.

It is important to point out that generation cannot be carried out without a lexicon containing the words. If different information pieces are annotated together as one single complex chunk, they need to be added in the lexicon as one single complex piece of information and it is necessary to process them ad hoc. The more information pieces are processed separately, the less ad hoc procedures have to be set.

⁹Figures like this one are taken from the video "La lepre e la tartaruga" [transl. "The hare and the tortoise"], in: Fiabe nel Bosco 1. DVD, Alba Cooperativa Sociale ONLUS, 2010.

4. The Annotation System

The fable has been annotated manually, glossing LIS sentences on different layers with Italian words.

We used this particular annotation method to allow generation and make the process as smooth as possible. Manual glosses are indeed a fairly free and subjective transcription system, since it is possible to annotate some features of signs and to omit other information, on the basis of what is considered relevant and appropriate to gloss.

The annotation has been organized in tables with eight simultaneous tiers dedicated respectively to affective (AFF), adverbial (ADV), syntactic (SYN) Non-Manual Markers, spatial agreement (AGR), Non-Manual Signs (NMS), Manual Signs (MS), Action Role Shift (ARS) and Quotation Role Shift (QRS). The first four tiers have been annotated in small letters, the remaining four in capitals.

The AFF tier is dedicated to affective NMMs, i.e. emotional states such as haughtiness, puzzlement, enthusiasm, surprise, embarrassment and so on.

The ADV layer conveys adverbial non-manual information, for example a particular facial expression realized to show that an action happens rapidly.

In the SYN tier syntactic NMMs are glossed, e.g. markers for topics and WH- and yes/no questions.

The AGR tier contains the annotation of specific points of the signing space used in the fable to establish thematic relations and anaphoric reference and that helped define arguments. Points of space indicated by pointing signs have also been annotated in this layer. These glosses consist of three parts: the prefix LOC that stands for "location", a number that identifies the location sign in order of realization in relation to the other ones and the corresponding spatial feature. Number and spatial features are separated by an underscore "_". More precisely, spatial features have been converted in this context into PROX (proximal), NPROX (non-proximal), DIST (distal), NDIST (non-distal). An example of annotated location is LOC1_NPROX.

Lexical NMMs are not included, because they were not relevant for our purpose and would have uselessly weighed down the annotation. Since the sign would be ungrammatical if such NMMs were not produced, we treated these expressions as they were incorporated covertly to the corresponding sign in the MS layer.

Furthermore, we did not take into account prosody head nods, fillers or hesitation markers used by the signer during the narration. However, we indicated pauses using punctuation (points, commas and colons) and inserted them in the MS tier, together with manual signs.

The NMS tier contains linguistic information that may be conveyed without the use of hands¹⁰ and it is closely related to Action Role Shift. NMS (Non-Manual Signs) seem to appear only under ARS and they seem to be body movements that acquire relevant linguistic meaning under RS

¹⁰A type of signs realized without the use of hands has already been introduced by (Dively, 2001) as Non Handed Signs (NHS). We intentionally did not use this terminology because we consider NMS and NHS to be part of two different categories.

and therefore have to be translated. An example of a NMS is the movement of the signer’s eyes – and head, eventually – towards the point of space in which the embodied interlocutor is located, that is translated in the fable as *GUARDARE*, which means “to look at”. “to pant”, annotated in the fable as *SBUFFARE*, can be the corresponding translation of the cheeks and mouth movement performed when we breath energetically and spasmodically through our mouth. As the reader may infer, two NMS may be simultaneously realized.

In NMS and MS tiers, dashes “-” are used to link words when the LIS sign or information conveyed can only be translated with more than one Italian word. Words linked with dashes are processed by the generator as one single term. Classifiers are usually marked with “_CL” at the end of the gloss. Annotations that end with “_L”/“_L” or “_T”/“_T” correspond to signs related to the hare (*lepre* in Italian) and the tortoise (*targaruga* in Italian), respectively. These suffixes are used to distinguish classifiers that indicate paws of either the tortoise or the hare or actions performed by the two animals and also indexicals that point to places in the space in which they are located. The symbol “+” at the end of a sign gloss indicates repetition. Simultaneous signs are glossed in round brackets “()”.

Pointing signs are annotated in the MS level. Signs referring to locations in the space are introduced by the abbreviation for “index” (IX) and their annotation structure follows the annotation structure of points in the space, e.g. IX-LOC1_NPROX.

The gloss IX-1 refers to the first-person singular pronoun. IX-3L and IX-3T indicate the hare and the tortoise respectively.

The ARS tier marks the use of the Action Role Shift related to the embodied character, such as the hare (glossed with L) and the tortoise (T).

The QRS tier includes words or thoughts of the characters and, in simple terms, defines discourse turns.

When necessary, underscores have been used to mark the onset of the realization of non-manual material and their occurrence. The offset can also be marked by the end of the corresponding word, as in (1):

- (1) SYN _____wh
 QRS [T_____]
 MS RUN-LAP HOW.
 “How is the race course like?”

The first sentence of the fable with English glosses is shown in Table 1 on the next page. Some information is called out with an asterisk due to lack of space in the table. Translation is also shown below.

5. Main Problems Encountered

As the other sign languages, LIS does not have a standardized and recognized writing system. Thus, the organization of glosses has been essential for us in order to make the generator generate an acceptable Italian text.

However, we had to deal with some problems during the annotation phase and the overall generation process.

5.1. Simultaneity

As for the annotation of the fable, our aim was to avoid creating a too restrictive and informative MS gloss that conveyed information that could be conveyed in other layers. Therefore, we organized gloss tiers trying to distribute different types of information on different levels. In this way, ad hoc procedures in the generation process would have been limited. In some cases, however, this was not possible. Let us take a practical example.

At the beginning of the fable, the signer tells that the hare is hopping, by embodying him. Then he produces the sign in Figure 3 below and narrates that the hare sees a tortoise.



Figure 3: “OCCHI_CL-VELOCE-VS-LOC2_NPROX”

In Figure 3, the picture left represents the beginning of the realization of the sign and the picture right is its end.

The sign in Figure 3 is a classifier: in this context, the signer’s hands represent the eyes of the hare, which move to his right, toward the position in which will then be located the tortoise. During the realization of this sign the movement of the hands is rapid and sharp and represents the movement performed by the eyes of the hare. The information conveyed is also adverbial: his eyes moved rapidly/sharply. This information is incorporated into the sign and it is conveyed manually by the movement parameter. Let us not forget that movement is an integral part of the sign. Therefore, the sign in Figure 3 has been glossed in the LIS Manual Signs layer as a single term: OCCHI_CL-VELOCE-VS-LOC2_NPROX, which would be EYES_CL-RAPID-TO-LOC2_NPROX glossed in English. This single sign gives us precise and detailed information: it alone could be translated as “(3ps) moved the eyes rapidly to the right”.

However, this is usually not the kind of information conveyed in a children’s fable in spoken languages, so in the target translation - written in order to evaluate the generated text in the thesis work - the information has been translated as follows: “While the hare was hopping, he suddenly saw a tortoise”.

Since the adverbial information is part of a single complex term in the MS layer in the glosses, it cannot be processed separately from the rest of the information contained in the same term. It cannot be extracted from it and cannot be identified by the generator as a single piece of information that is contained in the lexicon. Therefore, the whole term had to be processed ad hoc.

Hence, to put it briefly, simultaneity of SL is a problem for generation. It is not easy to transpose simultaneous SL information into different layers with specific functions that then have to be transformed in sequential strings. There is

* = LOC1_NPROX

AFF	
ADV	
SYN	__wh
AGR	_____*
NMS	
MS	WOOD, IX-LOC1_NPROX TO-LIVE WHO. HARE, HARE HAUGHTY.
ARS	
QRS	

“In the wood lived a hare, a haughty hare.”

Table 1: English glosses of the first sentence of the fable.

a tangible risk of creating too informative levels of glosses, which have to be processed ad hoc.

A possible solution could be to modify the organization of glosses and expand the MS layer, dividing it into four sub-layers that correspond to the phonological parameters: HandShape, MOVement, LOCation and Palm Orientation. When necessary, modifications to the standard realization of the sign may be indicated in these tiers, as shown in Table 2:

MS	EYES
HS	F
MOV	MOVE RAPIDLY
LOC	
PO	TO THE RIGHT

Table 2: “to move the eyes rapidly to the right”

MS layer contains information about what the sign refers to.

HS could be considered as a merely descriptive level that indicates modifications of the handshape of the generic sign. Figure 3, for example, depicts a classifier, i.e. a non-standard sign, which is realized with a different handshape compared to the handshape used in the realization of the generic sign for “eyes” in LIS.

MOV conveys the adverbial information given by the hand movement, here “rapidly”. In this case, it also tags the sign as verb.

LOC remains empty in this case, but it could be used for agreement purposes and anaphora reference, replacing the AGR layer created in the glosses above.

Finally, PO contains the meaning conveyed by changes in palm orientation, “to the right” in Table 2.

In this way the generation of all the information carried by the classifier should be possible. However this is a complex issue, which needs to be further investigated. Furthermore it should be checked if this subdivision could be used to organize information conveyed by other classifiers or signs too. Perhaps it is too restrictive or unsuitable.

5.2. LIS is Not (Signed) Italian

There are other signs involved in ad hoc procedures in the generation process of the fable and some of them are actually welcome problems, since they are also the proof that

LIS is not a basic system that matches “gestures” with the Italian grammar and structure. This is the case of “venne” (3ps “came”) instead of “arrivò” (3ps “arrived”) in the fable. In fact, we consider the sentence in (2) a little odd if used to narrate a fable:

- (2) Poi venne un gufo.
“Then came an owl.”

As Italian speakers, we prefer the use of “arrivò” (3ps “arrived”) instead. On the contrary, (3) in LIS is totally acceptable:

- (3) _____wh
POI VENIRE CHI. GUFO.
then to-come who? owl .
“Then comes an owl.”

5.3. No Base Tier in Glosses

It is also important to notice that there is not only one gloss tier that consistently conveys the relevant information. For example, it is not necessarily true that the MS layer is the most informative one, as we can see in Table 3:

* = LOC2_NPROX

AFF	
ADV	
SYN	
AGR	_____*
NMS	_____PANT _____LOOK-AT
MS	REFERENT1_CL IX-3L HARE, PAWS(L)_CL.
ARS	[L_____]
QRS	

Table 3: “The hare pants and looks to the right.”

In Table 3, the most part of the relevant information is conveyed through NMMs. In fact, PAWS(L)_CL is a Noun and Limb/Body Part Classifier, which helps us to understand that the hare is being embodied. However, the layers that explain what the hare does (Figure 4 on the next page) are NMS and AGR.

The same applies to Table 4 and Figure 5.

The generator receives as an input one-layer terms that contain slots, which correspond to the annotation tiers and can contain information or not. There is no base tier in glosses that consistently carries the relevant information or the in-



Figure 4: Embodiment of the hare panting and looking right

* = GETS-READY-TO-START

AFF	
ADV	
SYN	
AGR	
NMS	*
MS	TORTOISE [...] PAWS(T)_CL.
ARS	[T_____]
QRS	

Table 4: "The tortoise gets ready to start."



Figure 5: Embodiment of the tortoise getting ready to start the race

formation needed to generate acceptable sentences, so the generator may receive the relevant information from any of the slots. This obviously makes the generation process more complex.

5.4. Pragmatics

It is also worth mentioning that pragmatics and visual modality gave us a hard time.

As far as pragmatics is concerned, we would like to stress that many signs assume a particular meaning depending on context. For example, Figure 6 may represent both a tortoise that walks and a tortoise that starts walking, depending on context. This sign in LIS had been annotated as "to walk".

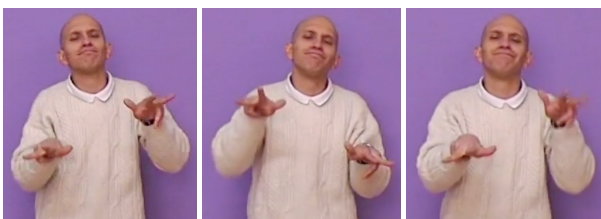


Figure 6: Embodiment of the tortoise walking

Italian allows using two different verbs to describe the two actions: "camminare" and "incamminarsi", respectively. The former is an activity; the latter implies an achievement and is an ingressive verb. It has been necessary to deal with the latter term on an ad hoc basis in order to generate it.

5.5. Visual Modality

Visual modality allows the immediate exchange of simultaneous information. When possible, signers take advantage of this modality using the space in front of them to locate entities and express their communicative intents. However, saying or explaining what we see as interlocutors is not always easy, let alone make a computational system generate it. In fact, some specific parts of the fable have not been generated correctly at the first attempt. For instance, this is the case of a part of the fable where the signer embodies the hare illustrating the route of the race to the tortoise. The illustrated route is depicted in Figure 7.

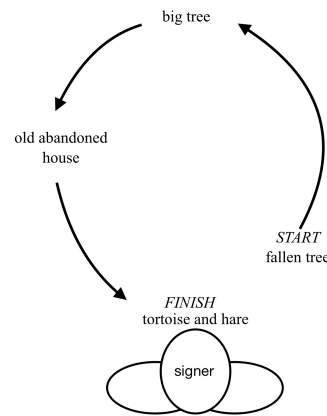


Figure 7: Route of the race in the fable

Segments covered by the arrows in Figure 7 are the parts of the space indicated by the signer using the index finger with a continuous movement to show the route of the race. These signs have been glossed in different ways in order to avoid redundancy in the generated text. In this scenario, the problem of redundancy does not present itself in LIS, even if the same sign is used for indicating. Actually, using that generic sign and keeping that handshape to indicate the entire route expresses cohesion. A change in handshape would probably imply a change in meaning. For example, if the signer embodying the hare used the sign in Figure 8 on the next page for a specific section of the route, he would probably want to specify to the tortoise that on that section they have to walk, instead of running.

However, the sign in Figure 8 would not be acceptable in this context, since it is usually used to refer to entities with two legs.

The correct generation of the sentences has been possible by glossing these continuous index signs as follows: INIZIARE-GIRO (START-LAP), CONTINUARE-GIRO (CONTINUE-LAP), FINIRE-GIRO (COMPLETE-LAP), respectively.

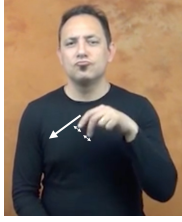


Figure 8: “to walk” (spreadthesign.com)

5.6. Verb Tenses and Moods

The previous example introduces another problem: verb tenses and moods.

The biggest obstacle was the generation of future tense, used in the fable to talk about the route of the race. In that context there are no signs that express future tense, therefore information about this tense could not be annotated. Nevertheless, who is watching the fable can easily infer that the mentioned actions are intended to take place in the future.

Unfortunately, it was not easy to find a well-defined linguistically motivated trigger to generate this tense. At first we thought it was only possible to infer it on a theoretical level or on a level of commonsense knowledge. Then we decided to focus on speech acts.

Let us give a practical example, where the hare is speaking:

- (4) IX-LOC3_PROX PARTIRE IX-LOC3_PROX.
 here to-start-from here .
 “Here, we start / are going to start from here.”

By imposing the illocutive¹¹ tag to the verb *partire* we have been able to produce *partiremo* (1pp of *partire*, future tense). However we could not use the same tag for INIZIARE-GIRO, CONTINUARE-GIRO, FINIRE-GIRO, so they appear in the generated text in present and not in future tense.

6. Conclusion

The paper presented the annotation system used to manually annotate the fable “The Tortoise and the Hare” signed in LIS.

LIS glosses have been the starting point of a generation process that involved the automatic generation of the fable in Italian. Glosses have been organized in eight simultaneous layers, i.e. affective, adverbial and syntactic Non-Manual Markers, spatial agreement, Non-Manual Signs, Manual Signs, Action Role Shift and Quotation Role Shift. In particular, the Non-Manual Signs tier was dedicated to linguistic information that may be conveyed without the use of hands and that seems to appear only under Action Role Shift.

We encountered some problems in the annotation and generation process and discussed them in this work. In particular, we had to deal with some peculiarities of SLs that are

¹¹An illocutive (or illocutionary) act is performed by the speaker by virtue of uttering certain words. Please see Austin (1962) e Searle (1969) for further study.

apparently not suitable for one-string generation, such as simultaneity. For example, hands alone in SLs can simultaneously convey different pieces of information - e.g. adverbs, verbs, etc. -, which have been glossed as single complex terms and therefore had to be processed ad hoc by the generator. From this we can infer that the Manual Signs layer in our glosses was too informative and carried pieces of information that could be contained in other layers. In order to solve this problem and facilitate the generation process, we considered the possibility of expanding the MS tier. The idea is to divide it into four sub-layers, corresponding to handshape, movement, location and palm orientation of the sign. By doing so, it should be easier to deal with signs such as classifiers used in the fable. However, this modification has not been implemented in our system yet. Moreover, we pointed out that some issues encountered are strictly related to pragmatics and the visual modality of sign languages. From the above, it is possible to draw some important conclusions. On the one hand, we have demonstrated that the Natural Language Generation field can manage some complexities of the Italian Sign Language, although generators can process information only if it is conveyed as a single input string. On the other hand, further studies on the organization and subdivision of layers in glosses are needed. It would be also very interesting to take into consideration for generation other annotation systems and compare their effectiveness to that of manual glosses.

7. Bibliographical References

- Aarons, D. and Morgan, R. Z. (2003). Classifier predicates and the creation of multiple perspectives in South African Sign Language. *Sign language studies*, 3(2):125–156.
- Ahlgren, I. (1990). Deictic pronouns in Swedish and Swedish Sign Language. In *Theoretical Issues in Sign Language Research*, volume 1, pages 167–174. University of Chicago Press.
- Aikhenvald, A. Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford University Press.
- Ajello, R. (1997). Lingue vocali, lingue dei segni e ‘l’illusion mimétique’. In *Schrtbthair a aninm n-ogaim*, pages 17–30. Pacini Editore.
- Allan, K. (1977). Classifiers. *Language*, 53:285–311.
- Austin, J. L. (1962). *How To Do Things with Words*. Oxford University Press.
- Bahan, B. and Petitto, L. (1980). *Aspects of rules for character establishment and reference in American Sign Language storytelling*. Unpublished manuscript. Salk Institute for Biological Studies, La Jolla, CA.
- Bahan, B. (1996). *Nonmanual Realization of Agreement in American Sign Language*. Ph.D. thesis, Boston University, Boston Mass.
- Baker, C. and Padden, C. (1978). Focusing on the non-manual components of ASL. In P. Siple, editor, *Understanding language through sign language research*, pages 27–57. Academic Press.
- Baker-Shenk, C. (1983). *A micro analysis of the nonmanual components of American Sign Language*. Ph.D. thesis, University of California. Berkeley, USA. Unpublished.

- Battison, R. (1978). *Lexical borrowing in American Sign Language*. ERIC.
- Benedicto, E. and Brentari, D. (2004). *Where Did All the Arguments Go? Argument-Changing Properties of Classifiers in ASL*. *Natural Language and Linguistics Theory*, 22(4), pages 743–810.
- Bertone, C. and Cardinaletti, A. (2009). *La grammatica dello spazio nella LIS*. Libreria Editrice Cafoscarina.
- Bertone, C. (2009). La grammatica dello spazio nella LIS. In *Alcuni capitoli della grammatica della LIS*, pages 79–99. Libreria Editrice Cafoscarina.
- Bertone, C. (2011). *Fondamenti di Grammatica della Lingua dei Segni Italiana*. FrancoAngeli.
- Branchini, C. and Donati, C. (2009). Relatively different Italian Sign Language relative clauses. *Correlatives cross-linguistically*, 1:157–191.
- Branchini, C. and Geraci, C. (2011). L'ordine dei costituenti in LIS: risultati preliminari. In *Grammatica, lessico e dimensioni di variazione nella LIS*, pages 113–126. Franco Angeli.
- Branchini, C. (2007). On relativization and clefting in Italian Sign Language. *Sign Language & Linguistics*, 10(2):201–212.
- Branchini, C. (2014). *On relativization and clefting. An analysis of Italian Sign Language*. Berlin, Mouton De Gruyter.
- Caselli, M. C., Maragna, S., and Volterra, V. (2006). *Linguaggio e Sordità: Gestì, Segni e Parole nello Sviluppo e nell'Educazione*. Il Mulino.
- Cecchetto, C., Geraci, C., and Zucchi, S. (2009). Another way to mark syntactic dependencies: The case for right-peripheral specifiers in sign languages. *Language*, 85:278–320.
- Corazza, S. (1990). The Morphology of Classifier Handshapes in Italian Sign Language (LIS). *Sign Language Research: Theoretical Issues*, pages 71–82.
- Corina, D. P., Bellugi, U., and Reilly, J. (1999). Neuropsychological Studies of Linguistic and Affective Facial Expressions in Deaf Signers. *Language and Speech*, 42(2-3):307–331.
- Cormier, K., Fenlon, J., Rentelis, R., and A., S. (2011). Lexical frequency in British Sign Language conversation: a corpus-based approach. In *Proceedings of the Conference on Language Documentation and Linguistic Theory 3*, pages 81–90. School of Oriental and African Studies.
- Craig, C. (1992). Classifiers in a functional perspective. In *Layered structure and reference in a functional perspective*, pages 277–301. John Benjamins.
- Cuxac, C. (2000). Iconicity of Sign Language. In *The structure of multimodal dialogue*, pages 321–334. Benjamins.
- Delmonte, R. and Trolvi, S. (2020). *Generating Text from Glosses in LIS*. Not Published.
- Dively, V. (2001). Signs without hands: nonhanded signs in American Sign Language. *Signed languages: Discoveries from international research*, pages 62–73.
- Donati, C., Barberà, G., Branchini, C., Cecchetto, C., and Quer, J. (2017). Searching for imperatives in European sign languages. *Imperatives and Directive Strategies*, pages 111–155.
- Emmorey, K. (1999). Do signers gesture? In Lynn S. Messing et al., editors, *Gesture, speech, and sign*, pages 133–159. Oxford University Press.
- Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: The Semantics and Morphosyntax of the Use of Space in a Visual Language*. Signum.
- Engberg-Pedersen, E. (1995). Point of view expressed through shifters. In *Language, Gesture, and Space*, pages 133–154. Lawrence Erlbaum.
- Franchi, M. L. (2004). Componenti non manuali. In V. Volterra, editor, *La lingua dei segni italiana*, pages 159–177. Il Mulino.
- Geraci, C., Cecchetto, C., and Zucchi, S. (2008). Sentential complementation in Italian Sign Language. In M. Grosvald et al., editors, *Western Conference On Linguistics (WECOL)*, pages 46–58, Davis, Department of Linguistics, University of California at Davis.
- Geraci, C. (2002). *L'ordine delle parole nella LIS*. Ph.D. thesis, Università Statale di Milano.
- Geraci, C. (2006). Negation in LIS (Italian Sign Language). In *Proceedings of NELS*, volume 35, pages 217–229. Batherman LC, Ussery.
- Herrmann, A. and Pendzich, N.-K. (2018). Between narrator and protagonist in fables of German Sign Language. In *Linguistic foundation of narration in spoken and sign languages*, pages 275–308. ohn Benjamins Publishing Company.
- Herrmann, A. and Steinbach, M. (2012). Quotation in Sign Languages. A Visible Context Shift. In *Quotatives. Cross-linguistic and Cross-disciplinary Perspectives*, pages 203–228. Benjamins.
- Herrmann, A. (2013). *Modal and focus particles in sign languages: A cross-linguistic study*, volume 2. Walter de Gruyter.
- Hickok, G., Bellugi, U., and Klima, E. S. (1996). The neurobiology of sign language and its implications for the neural basis of language. *Nature*, 381(6584):699.
- Klima, E. S. and Bellugi, U. (1979). *The Signs of Language*. Harvard University Press.
- Lee, R. G., Neidle, C., MacLaughlin, D., Bahan, B., and Kegl, J. (1997). Role shift in ASL: A syntactic look at direct speech. In *Syntactic Structure and Discourse Function: An Examination of Two Constructions in American Sign Language*. American Sign Language Linguistic Research Project Report No. 4.
- Liddell, S. K. (1978). Non-manual signals and relative clauses in American Sign Language. In P. Siple, editor, *Understanding language through sign language research*, pages 59–90. Academic Press.
- Liddell, S. K. (1980). *American Sign Language Syntax*. The Hague: Mouton De Gruyter.
- Liddell, S. K., (1995). *Real, Surrogate, and Token Space: grammatical Consequences in ASL*, pages 19–41. Lawrence Erlbaum.
- Liddell, S. K., (2002). *Indicating verbs and Pronouns, Pointing away from Agreement*, pages 303–320. Lawrence Erlbaum.

- Lillo-Martin, D. (1995). The Point of View Predicate in American Sign Language. In *Language, Gesture, and Space*, pages 155–170. Lawrence Erlbaum.
- Lillo-Martin, D. (2012). Utterance reports and constructed action. In *Sign Language – An International Handbook*, pages 365–387. Mouton de Gruyter.
- Mac Laughlin, D. (1997). *The Structure of Determiner Phrase: Evidence From American Sign Language*. Ph.D. thesis, Boston University.
- Mazzoni, L. (2008). Classificatore del corpo e impersonamento in LIS. In *I segni parlano. Prospettive di ricerca sulla Lingua dei Segni Italiana*, pages 64–75. Francoangeli.
- Mazzoni, L. (2009). Evidenzialità e Impersonamento in LIS. In *La grammatica della Lingua dei Segni Italiana. Atti dell'incontro di studio*, pages 79–99. Libreria Editrice Cafoscarina.
- Mazzoni, L. (2012). *Classificatori e impersonamento nella Lingua dei Segni Italiana*. Pisa University Press.
- McCullough, S., Emmorey, K., and M, S. (2005). Neural organization for recognition of grammatical and emotional facial expressions in deaf ASL signers and hearing nonsigners. *Cognitive Brain Research*, 22(2):193–203.
- Meier, R. P. (1990). Person Deixis in American Sign Language. In *Theoretical Issue in Sign Language Research*, volume 1, pages 175–190. University of Chicago Press.
- Metzger, M. (1995). Constructed dialogue and constructed action in American Sign Language. In C. Lucas, editor, *Sociolinguistics in Deaf Communities*, pages 255–271. Gallaudet University Press.
- Padden, C. (1986). Verbs and role shifting in ASL. In C. Padden, editor, *Proceedings of the Fourth National Symposium on Sign Language Research and Training*, pages 44–57. Silver Spring MD, National Association of Deaf.
- Padden, C. (1990). The relation between Space and Grammar in ASL Verb Morphology. In C. Lucas, editor, *Sign Language Research, Theoretical Issues*, pages 118–132. Gallaudet University Press.
- Pfau, R. and Quer, J. (2010). Nonmanuals: their grammatical and prosodic roles. In Diane Brentari, editor, *Sign Languages*, Cambridge Language Surveys, pages 381–402. Cambridge University Press.
- Pizzuto, E. and Corazza, S. (1996). Noun morphology in Italian Sign Language (LIS). *Lingua* 98, pages 169–196.
- Pizzuto, E., Giuranna, E., and Gambino, G., (1990). *Manual and nonmanual morphology in Italian Sign Language: grammatical constraints and discourse processes*, pages 83–102. Gallaudet University Press.
- Pizzuto, E., Cameracanna, E., Corazza, S., and Volterra, V., (1995). *Terms for spatio-temporal relations in Italian Sign Language*, pages 237–256. John Benjamins.
- Quer, J. (2016). Reporting with and without role shift: sign language strategies of complementation. In *A Matter of Complexity: Subordination in Sign Languages*, pages 204–230. De Gruyter Mouton.
- Schembri, A. (2003). Rethinking ‘Classifiers’ in Signed Languages. In K. Emmorey, editor, *Perspectives on Classifiers in Sign Language*, pages 3–34. Lawrence Erlbaum Associates.
- Schlenker, P. (2017a). Super monsters I: Attitude and action role shift in sign language. *Semantics and Pragmatics*, 10(9):1–65. Early access version.
- Schlenker, P. (2017b). Super monsters II: Role Shift, iconicity and quotation in sign language. *Semantics and Pragmatics*, 10(2):1–67. Early access version.
- Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.
- Stokoe, W. (1960). Sign language structure: An outline of the visual communication systems of the American Deaf. *Studies in Linguistics: Occasional Papers*.
- Tomaszewski, P. and Farris, M. (2010). Not by the Hands Alone: Functions of Non-Manual Features in Polish Sign Language. In B. Bokus, editor, *Studies in the psychology of language and communication*, pages 289–320. Matrix.
- Valli, C. and Lucas, C. (2000). *Linguistics of American sign language: an introduction*. Gallaudet University Press.
- Volterra, V., Laudanna, A., Corazza, S., Radutzky, E., and Natale, F. (1984). Italian Sign Language: the order of elements in the declarative sentence. *Recent research on European Sign Languages*, pages 19–48.
- Volterra, V., Laudanna, A., Corazza, S., Radutzky, E., and Natale, F. (1986). The verb system of Italian Sign Language. *Signs of life*, pages 17–31.
- Volterra, V. (2004). *La lingua dei segni italiana: La comunicazione visivo-gestuale dei sordi*. Il Mulino.
- Wilbur, R. B. (2000). Phonological and Prosodic Layering of Nonmanuals in American Sign Language. In K. Emmorey et al., editors, *The Signs of Language Revisited: an anthology to honor Ursula Bellugi and Edward Klima*, pages 215–244. Lawrence Erlbaum Associates.
- Woodward, J. C. (1972). Implications for Sociolinguistic Research among the Deaf. *Sign Language Studies*, 1(1):1–7.
- Zucchi, A. (2004). *Monsters in the visual mode*. Unpublished Manuscript. Università degli Studi di Milano.
- Zucchi, S. (2009). Along the Time Line: Tense and Time Adverbs in Italian Sign Language. *Natural Language Semantics*, 17:99–139.