



Università  
Ca'Foscari  
Venezia

Master's Degree  
in Scienze del Linguaggio

Final Thesis

# Generating Italian from Italian Sign Language glosses with GENLIS

**Supervisor**

Ch. Prof. Marina Buzzoni

**Assistant supervisor**

Ch. Prof. Rodolfo Delmonte

**Graduand**

Serena Trolvi

Matriculation number 841996

**Academic Year**

2018 / 2019



*Alla mia famiglia*



# Abstract in deutscher Sprache

Gebärdensprachen sind die natürlichen Sprachen tauber Menschen.

Im Gegensatz zur allgemeinen Überzeugung sind sie eigenständige und vollwertige Sprachen, die sich von Lautsprachen unterscheiden. Außerdem sind sie keine bloße Gestensysteme und keine Pantomime. Sie sind auch nicht international: Jedes Land hat seine eigene Gebärdensprache und manche Ländern haben mehr als eine (Caselli *et al.* 2006).

Dies gilt auch für die italienische Gebärdensprache (LIS), die die in Italien benutzte Gebärdensprache ist, die aber auf nationaler Ebene noch nicht anerkannt ist. Wie die anderen Gebärdensprachen ist sie eine visuell-manuelle Sprache, die eigene Grammatik, Phonologie, Morphologie, Syntax und Wortschatz hat (vgl. u.a. Volterra 1987 [2004]; Caselli *et al.* 2006; Geraci 2006; Geraci *et al.* 2008; Cecchetto *et al.* 2009; Branchini und Geraci 2010). Gebärdensprachen haben daher spezifische Eigenschaften und bedienen sich ihrer eigenen Komponenten, die manuell oder non-manuell sind, um Informationen gleichzeitig zu übertragen. Manuelle Komponenten sind Gebärden und bestehen aus vier Parametern:

- i. Handform, die die äußere Gestalt der Hand ist;
- ii. Handstellung, d.h. die Ausrichtung der Hand;
- iii. Ausführungsstelle, die die Stelle am Körper oder im Gebärdenraum<sup>1</sup> ist, wo die Gebärde vollzogen wird;
- iv. Bewegung, die die Bewegung der Hand beschreibt, d.h. wie und wohin die Hand bewegt wird. (vgl. zur amerikanischen Gebärdensprache, Stokoe 1960 und Battison 1978; zu LIS, Volterra [1987] 2004).

Non-manuelle Komponenten umfassen hingegen den Gesichtsausdruck, die Bewegung des Oberkörpers, des Kopfes, der Augenbrauen, der Wangen, des Mundes und der Schultern und die Blickrichtung. Außerdem können sie linguistische und non-linguistische Funktionen haben (vgl. u.a. Corina *et al.* 1999; McCullough *et al.* 2005; Pfau und Quer 2010; Herrmann 2013) und einigen Experten zufolge zählen sie zu dem fünften der oben genannten Parameter (Valli und Lucas 2000).

---

<sup>1</sup> Der Gebärdenraum ist der Raum vor dem Körper, in dem Gebärden ausgeführt werden.

In Anbetracht dieser Vorbemerkung, ist es nicht schwierig zu begreifen, dass die Simultaneität eine Kerneigenschaft der Gebärdensprachen ist. Diese taucht auch bei der Verwendung des Gebärdenraums, des sogenannten *Role Shift*<sup>2</sup> und bestimmter Gebärden wie Klassifikatoren<sup>3</sup> auf.

Besonders interessant wäre es deshalb herauszufinden, ob und inwiefern einem Computer, der Informationen nur sequentiell verarbeiten kann, es irgendwie gelingt, eine Gebärdensprache zu verarbeiten.

Diese Überlegungen haben mich dazu geführt, die automatische Generierung eines italienischen Textes aus Glossen von einer Fabel auf Italienische Gebärdensprache zum Gegenstand der vorliegenden Arbeit zu machen.

Bevor die dabei verwendeten Methoden beleuchtet werden, ist es wichtig in diesem Zusammenhang den Begriff „Generierung“ zu definieren. Mit „natürlicher Sprachgenerierung“ (NLG – *Natural Language Generation*) bezeichnet man die automatische Produktion von Texten in natürlicher Sprache (Output) aus nicht-sprachlicher Information (Input) durch eine maschinelle Verarbeitung bzw. einen Generator Algorithmus. Die Kernaufgabe eines solchen NLG-Systems ist es, Wahlen zu treffen (Hovy 1988; Reiter 2010). Es muss nämlich entscheiden, was es generieren soll und wie es dies tun kann. Hierfür muss es bestimmte Aufgaben ausführen, die in dieser Arbeit erklärt werden. Zahlreich sind natürlich die Anwendungsgebiete für automatische Sprachgenerierung, wie zum Beispiel die Generierung von Wetterberichten (Goldberg *et al.* 1994) und von geschriebenen oder mündlichen Texten, die Behinderten helfen, bestimmte Aufgaben zu erfüllen (Ferres *et al.* 2006; Reiter *et al.* 2009).

Diese Arbeit befasst sich mit der Generierung der Fabel „Die Schildkröte und der Hase“, die in bestimmten Schritten ausgeführt wurde.

Hierbei wurde zunächst die Fabel durch Glossentranskription verschriftlicht. Diese ist ein sehr subjektives Notationssystem, das es ermöglicht, die Sprache schriftlich abzubilden. Sätze auf LIS wurden nämlich mit italienischen Worten so notiert, dass die

---

<sup>2</sup> *Role Shift* ist eine Narrationsstrategie, die dem Gebärdenden es ermöglicht, etwas aus der Perspektive einer anderen Person oder eines Gegenstandes zu erzählen.

<sup>3</sup> Klassifikatoren sind im Wesentlichen grammatische Einheiten, die in einigen Sprachen verwendet werden, um Substantive zu klassifizieren.

von ihnen vermittelten Informationen auf acht parallel laufende Schichten (AFF, ADV, SYN, AGR, NMS, MS, ARS, QRS) verteilt wurden. In der AFF Schicht werden beispielsweise non-manuelle Komponenten transkribiert, die Emotionen ausdrücken; SYN enthält non-manuelle Komponenten, die eine syntaktische Funktion haben; in MS werden manuelle Komponenten bzw. Gebärden notiert, usw. .

Des Weiteren wurden die acht Schichten in Strings konvertiert, damit Glossen als Input für unseren Generator fungieren konnten. Darüber hinaus wurden Glossen in semantische Formen umgewandelt. Dies ist der Ausgangspunkt, von dem aus unserer Generator GENLIS seine Aufgabe erledigen kann.

In dieser Arbeit wird außerdem GENLIS vorgestellt, der auf bestimmten Algorithmen basiert. Der Kernalgorithmus ist ein Grammatik-Algorithmus, der Nebenregeln verwendet, die notwendig sind, um stilistisch markierte Strukturen zu generieren. Eingeführt werden auch weitere Algorithmen, die dem Generator es ermöglichen, den Output zu generieren.

Wie schon erwähnt, ist der von uns gewählte Text eine Fabel. Fabeln sind Erzählungen mit belehrender Absicht und sind normalerweise an Kinder gerichtet. Auf Italienisch werden sie grundsätzlich in der Vergangenheit erzählt. Dies gilt für LIS allerdings nicht: Gebärdender erzählen eher eine Geschichte in ihrem Hier und Jetzt, indem sie in der Geschichte ihre Raum-Zeit-Koordinaten verwenden, d.h. sie erzählen, als ob das Ereignis in ihrer Gegenwart passieren würde. Als Kennzeichen des LIS-Erzählens gelten zudem bestimmte linguistische Strategien und Elemente, wie die oben genannten Klassifikatoren und *Role Shift*, die in italienischen Fabeln nicht zu finden sind. Solche Faktoren haben wir natürlich berücksichtigt.

Weiterhin wird der generierte Text in dieser Arbeit vorgestellt und analysiert. Insbesondere wird er mit dem von uns verfassten Zieltext verglichen. Dieser diente als Muster, um den Output zu bewerten und beurteilen. Wir haben ihn abgefasst, indem wir versucht haben, uns nicht zu weit von dem Ausgangstext zu entfernen. Wir haben auch in Betracht gezogen, was der Generator eigentlich generieren könnte. Infolgedessen haben wir einen Text geschrieben, der auch Verben in *passato remoto* und *imperfetto* (Präteritum) enthält und alles in allem korrekt und wohlgeformt ist, obwohl einige Sätze scheinen, die Grenzen der Akzeptabilität ein wenig zu überschreiten.

Im Großen und Ganzen sind wir mit dem von GENLIS generierten Text zufrieden. Er ist leserlich, Informationen sind in logischer Reihenfolge geordnet und er ist ziemlich ausführlich im Vergleich zu Glossen. Im Text tauchen jedoch Probleme auf und die vorliegende Arbeit setzt sich auch damit auseinander, diese Schwierigkeiten zu erkennen. Es wird erläutert, dass der generierte Text an einigen Stellen starr, monoton und redundant ist. Es gibt auch verschiedene Fehler, die beispielsweise mit Kongruenz, direkter Rede und anaphorischen Verbindungen verbunden sind. Darüber hinaus wurden Tempora oft falsch generiert.

Im Laufe dieser Masterarbeit wurde bemerkt, dass einige von den oben genannten Problemen durch die Weise verursacht sind, in der Glossen transkribiert wurden. Dies ist darauf zurückzuführen, dass eine manuelle Gebärde alleine mehrere Informationen gleichzeitig weitergeben kann, die oft kontextabhängig sind. Eine solche Gebärde würde (und wurde) auf Italienisch mit einer Periphrase übersetzt. Daraus folgt, dass die annotierte MS Schicht manchmal zu informativ ist, indem sie Signifikate enthält, die eine andere mögliche Schicht enthalten könnte (z.B. manuelle adverbiale Informationen). Die Generierung aus solchen Gebärdenglossen musste deshalb ad hoc vollzogen werden. Aus diesem Grund werden am Ende dieser Arbeit einige Änderungen an Glossen vorgeschlagen.

Wie oben betont, kann eine Gebärde in vier Parameter zerlegt werden, die unterschiedliche Informationen übertragen. Veranschaulicht wird deshalb die Untergliederung der MS Schicht in vier untergeordnete Schichten, jede von denen einem Parameter entspricht und ein Stück Information überträgt. Wenn auch diese als Input verarbeitet würden, sollte es einfacher sein, einen korrekten Output zu generieren. Es muss auf jeden Fall geprüft werden, ob diese Untergliederung eine für die Generierung geeignete Strategie ist.

Abschließend wird in dieser Masterarbeit nachgewiesen, dass LIS eine echte Sprache ist und dass es notwendig ist, weitere Forschung in diesem Gebiet durchzuführen. Nebenher hat es sich deutlich ergeben, dass es möglich ist, einen italienischen Text aus LIS Glossen zu generieren, obwohl vielerlei Probleme während des Generierungsprozesses auftauchen können. In unserem Fall sollen die Glossentranskriptionsmethode und GENLIS selbst in der Zukunft optimiert werden.





# Contents

Annotation Conventions . . . . .	14
----------------------------------	----

<b>Introduction . . . . .</b>	<b>15</b>
-------------------------------	-----------

## **Chapter 1. The Italian Sign Language – Theoretical Framework**

1.1. Introduction . . . . .	17
1.2. Background . . . . .	18
1.3. Phonology . . . . .	21
1.4. Morphology. . . . .	24
1.4.1 Nouns and verbs . . . . .	24
1.4.1.1 Noun morphology . . . . .	26
1.4.1.2 Verbal morphology . . . . .	28
1.4.2 The use of space . . . . .	30
1.4.3 Time, tense and aspect . . . . .	33
1.4.4 Classifiers . . . . .	35
1.5. Syntax . . . . .	40
1.5.1 Word order . . . . .	41
1.5.2 Other general properties . . . . .	42
1.6. Non-Manual Markers . . . . .	44
1.7. Role Shift . . . . .	50
1.8. Writing systems for Sign Language . . . . .	53
1.9. Summary . . . . .	58

## **Chapter 2. Processing a Natural Language**

2.1	Introduction . . . . .	59
2.2	Natural Language Generation . . . . .	60
2.2.1	What is NLG? . . . . .	60
2.2.2	The big questions for NLG . . . . .	62
2.2.3	NLG tasks . . . . .	63
2.2.4	NLG approaches . . . . .	68
2.2.5	NLG applications . . . . .	70
2.2.6	Some examples of NLG systems . . . . .	72
2.3	From words to signs, from signs to words . . . . .	77
2.3.1	Foreword . . . . .	77
2.3.2	From vocal to sign language . . . . .	80
2.3.3	From sign to vocal language . . . . .	85
2.4	Summary . . . . .	89

## **Chapter 3. The Experiment**

3.1	Introduction . . . . .	91
3.2	Fables . . . . .	92
3.3	“The Tortoise and the Hare”. . . . .	94
3.3.1	The video . . . . .	94
3.3.2	Annotation . . . . .	95
3.3.2.1	Annotation tiers . . . . .	95
3.3.3	The target Italian version . . . . .	99
3.3.4	The generated text. . . . .	101
3.4	Summary . . . . .	106

## **Chapter 4. The GENLIS Generator**

4.1	Introduction . . . . .	107
4.2	Generation mechanism . . . . .	108
4.2.1	Manual glosses and semantic forms in Prolog . . . . .	109
4.2.2	The Generation Algorithm . . . . .	112
4.2.3	The Algorithm for Definiteness Assignment . . . . .	118
4.2.4	The Algorithm for Narrative direct speech speaking verb type . . . . .	121
4.2.5	Mapping Tense and Mood from Speech Act and Verbal Aspectual Lexical Properties . . . . .	123
4.3	Main problems . . . . .	125
4.4	Summary . . . . .	132

<b>Conclusions</b> . . . . .	135
------------------------------	-----

Appendix A . . . . .	139
----------------------	-----

Appendix B . . . . .	169
----------------------	-----

Appendix C . . . . .	172
----------------------	-----

References . . . . .	177
----------------------	-----

Webography . . . . .	197
----------------------	-----

## Annotation conventions

GLOSS	Gloss identifying a manual sign
GLOSS-GLOSS	Two or more English words corresponding to one sign
IX	(generic) pointing sign
IX-1	Pointing sign functioning as personal pronoun
<u>NMMs</u>	Non-manual marking (above) simultaneously co-
GLOSS	occurring with the manual sign (under)
*	Judgment of ungrammaticality

# Introduction

Sign Languages (SLs) are languages used by Deaf<sup>1</sup> communities. In contrast to what many people still think, they have their own grammar and specific phonological, morphological and syntactic rules. The same applies to Italian Sign Languages (LIS), which is the sign language used by the Italian Deaf community, although it lacks official recognition at national level to date.

In other words, SLs are natural languages, like vocal languages. However, there is a crucial difference between them: for vocal languages communication occurs in the vocal-auditory channel; 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 hands: non-manual articulators like head, eye gaze and shoulders play a crucial role too. Hence, it can easily be inferred that information is often conveyed simultaneously. However, hands alone can express several pieces of simultaneous information too.

Simultaneity is a property of grammar of SLs and SLs make extensive use of it. A couple of years ago I started wondering if such languages could be processed by a computer, which elaborates information sequentially. In the present thesis LIS meets Computational Linguistics, and more precisely Natural Language Generation (NLG). An attempt is made to generate written Italian from glosses of the fable “The Tortoise and the Hare” signed in Italian Sign Language. I use the term *generate* and not the term *translate* intentionally. I will explain why in this work.

This thesis is divided into four chapters.

Chapter 1 presents the theoretical framework of LIS. I will first provide the reader some background information, by dispelling several misconceptions about (Italian) Sign Language. This also allows demonstrating that SLs are “real” languages, and not some sort of gestural system with no linguistic status. Then, the chapter concentrates on describing LIS from a phonological, morphological and syntactic point of view. A section will also be devoted to Non-Manual Markers and to the phenomenon of Role

---

<sup>1</sup> Capitalized letter in the term “Deaf” refers to deaf people who identify themselves as members of a signing community. The word “deaf” with lower case letter indicates only the condition of hearing loss.

Shift. Finally, I will mention the main writing systems used to transpose sign languages into written form.

Chapter 2 gives an overview of the Natural Language Generation field. More precisely, it focuses on explaining what NLG is and what its tasks and applications are. The reader will also be introduced to some NLG systems that generate vocal and sign languages.

In Chapter 3 the focus shifts to the main subject of the work, which is our experiment on the generation of Italian starting from Italian Sign Language glosses. The reader will be first introduced to the domain of fables. I dedicate the next sections to the fable “The Tortoise and the Hare”: I will present the LIS video and describe the annotation methodology I utilized. I will then illustrate our Italian translation, which was used as target translation. This will be compared to the output text of the generator, which will be illustrated and discussed in the last section of the chapter.

Finally, Chapter 4 introduces GENLIS, the machine that generates from Italian Sign Languages glosses to Italian. The chapter is organised in two main sections: in the former the generation mechanism is shown; in the latter the main problems encountered in the generation are listed.

# Chapter 1

## The Italian Sign Language Theoretical Framework

### 1.1 Introduction

Sign Languages are natural human languages that convey meaning through the visual-spatial modality. They are equal to spoken languages at all linguistic levels, since they have their own grammar, structure and meaning.

The Italian Sign Language (LIS), which has not received an official linguistic recognition at national level yet, is the Sign Language used by the Italian Deaf community.

In the present chapter I will focus on the linguistic status of LIS: I will debunk some general false misconceptions on sign languages and I will demonstrate that they, and more specifically LIS, possess their own phonology, morphology and syntax. Furthermore, the reader will be introduced to some peculiarities of these languages, such as Non-Manual Markers and the use of Role Shift. Finally, the last section will be dedicated to the main writing systems for sign languages.

## 1.2 Background

Sign Languages are languages used by members of Deaf communities.

Academic interest for sign languages arose in the late 1950s with people like William Stokoe, an American researcher who published the first linguistic analysis of American Sign Language (ASL) and proposed that ASL is a fully formed human language, equal to a spoken language. He demonstrated that Sign Languages are not mere gestural systems without grammar.

As for Italian Sign Language (LIS), studies began some years later, in the late '80s, thanks to Virginia Volterra and her team of researchers.

Despite the countless works on the linguistic status of sign languages (SLs), there are still many misconceptions around SLs. They will be dispelled in the next lines.

First of all, I would like to stress that sign languages are not universal. People not familiar with the Deaf world often ask me why all deaf people in the world do not use a unique sign language. “It would be easier”, they say. The point is that SLs are natural languages, just like spoken languages. Each country generally has its own native sign language, and some countries have more than one. According to the *Ethnologue of world languages*<sup>1</sup> 144 sign languages exist to date. Every Deaf community uses its own language with its own peculiarities and the same sign can be used to express different meanings in two different countries or even in the same country. For instance, Figure 1, Figure 2 and Figure 3 show the same sign, which means “name” in ASL, “Rome” in LIS and “sick” in a variant of LIS used in the city of Trieste in northeastern Italy.



**Figure 1.** “Rome” (standard LIS)  
(Caselli *et al.* 2006:38)



**Figure 2.** “name” (ASL)  
(Caselli *et al.* 2006:38)



**Figure 3.** “sick” (variant of LIS of Trieste)  
(Russo Cardona e Volterra 2007:57)

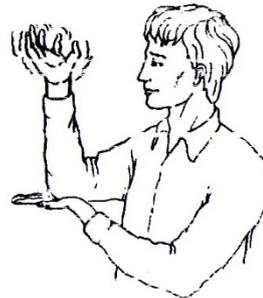
---

<sup>1</sup> [www.ethnologue.com](http://www.ethnologue.com) (last consulted 9/7/2019).

However, the same meaning is usually expressed with different signs in different Sign Languages:



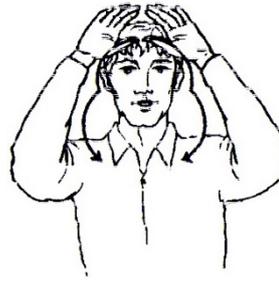
**Figure 4.** “tree” in LIS  
(Caselli *et al.* 2006:39)



**Figure 5.** “tree” in LSE<sup>2</sup>  
(Caselli *et al.* 2006:39)



**Figure 6.** “tree” in ASL  
(Caselli *et al.* 2006:39)



**Figure 7.** “tree” in AUSLAN<sup>3</sup>  
(Caselli *et al.* 2006:39)

Thus, we can claim that signs are conventional and arbitrary.

Nevertheless, I want to point out that an international Sign “Language”, known also as *Gestuno*, does exist. It is a pidgin sign language created to allow communication between signers from different countries. Actually, it is not a fully-formed language, but rather a vocabulary of “naturally spontaneous and easy signs in common use by deaf people of different countries” (British Deaf Association 1975:2).

Furthermore, it is a widespread belief that sign languages derive completely from spoken languages. I will prove that this is not true.

Let us first make a comparison between two vocal languages. We all know that American and British English are very similar. Therefore, one may believe that

---

<sup>2</sup> Spanish Sign Language.

<sup>3</sup> Australian Sign Language.

American and British Sign Languages are very similar too, but this is not the case. ASL and French Sign Language (LSF) are actually more intelligible than ASL and British Sign Language (BSL), due to historical reasons.

Moreover, if SLs were dependent from spoken languages, they should have the same syntactic structure, but they have not. Let us consider, for instance, the case of Italian and LIS. Italian is a SVO language and the WH-element generally appears at the beginning of the sentence. LIS instead seems to be a SOV language and LIS signers usually realize the WH-element at the end of their sentence. Anyway, from a sociolinguistic point of view it should be also mentioned that the fact that sign languages are not dependent from vocal languages does not mean that they cannot be influenced by them.

For the sake of completeness, I want also to clarify that there are signed communication systems that strive to be a representation of grammar and vocabulary of the respective vocal language; for example, as for Italian, Signed Italian and Signed Exact Italian. The first one uses vocabulary of LIS and the Italian grammatical structure. The second one applies the sign of Italian Sign Language to Italian word order and grammar and utilizes in addition fingerspelling and specific signs to refer to those vocal morphological elements that appear overtly in spoken languages, such as articles.

Another common misconception is that signs are completely iconic, that is they visually look like the definition they are trying to convey. Again, this is not true. They cannot be defined as gestures, because signs have an own complex inner structure and for this reason we cannot compare sign languages to non-linguistic systems like pantomime.

A first study on the comparison between ASL and pantomime comes from Klima and Bellugi (1979). They show that the realization of pantomimes is very different from one individual to another and it respects duration and directions of the movement of the represented event. In addition, pantomimes can be understood by anyone.

This does not apply to signs, because these are produced similarly in the same community and convey information about the event through simplified and restrained movements. It is also important to notice that the meaning of most signs is not easily understandable for people who don't know the sign language. However, some iconic signs do exist and signs may be more transparent than words (Pietrandrea 2000, 2002),

but transparency doesn't always seem to allow sign comprehension.<sup>4</sup> Therefore signs are used as they were arbitrary.

At this point, we can conclude that SLs are real languages with own linguistic systems. In fact, they have a phonology, morphology, syntax, semantics and pragmatics. I will focus on the first three of them in this chapter.

### 1.3 Phonology

Phonology is the branch of linguistics that deals with the organization of sounds within a language. In spoken languages, a unit of sound that has no meaning and distinguishes one word from another is called a phoneme. The easiest way to identify it is the minimal pair, which is a pair of words with different meaning that differ in only one phonological element (e.g. /b/ and /c/ in *bat* and *cat*). A phoneme in sign language is called a chereme, which is the smallest unit in which a sign can be represented.

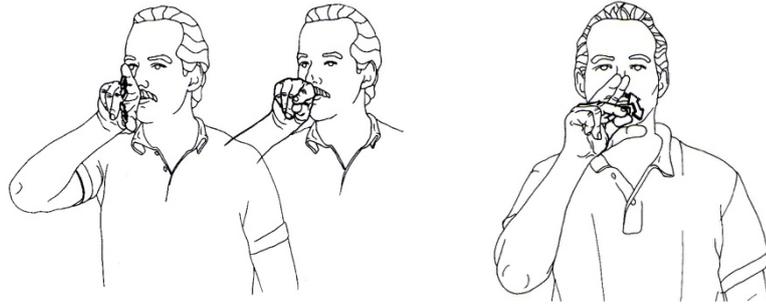
According to Stokoe (1960), a sign can be analyzed into three specific cheremes, or formational parameters: the handshape, which is the hand configuration; the location in which the signs are produced; and the movement, which describes the hand movement during the realization of the sign. Later on, Battison (1978) added a fourth parameter: the hand (or palm) orientation, which refers to the direction in which the hand is turned to produce the sign. In LIS 38 handshapes, 15 locations, 32 movements and 6 palm orientations have been identified (Volterra [1987] 2004).

More recent studies (Valli and Lucas 2000) have also demonstrated the existence of a fifth parameter of sign languages: Non-Manual Markers (NMMs), which include facial expressions and head and body movements. This topic will be discussed in detail in Section 1.6.

The examples on the next page show some minimal pairs in LIS:

---

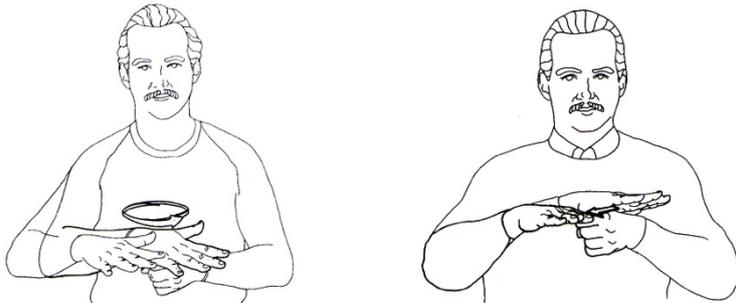
<sup>4</sup> For further information about this topic see Frishberg (1975) and Klima and Bellugi (1979) for ASL, Grosso (1992-1993, 1997) and Caselli *et al.* (2006) for LIS.



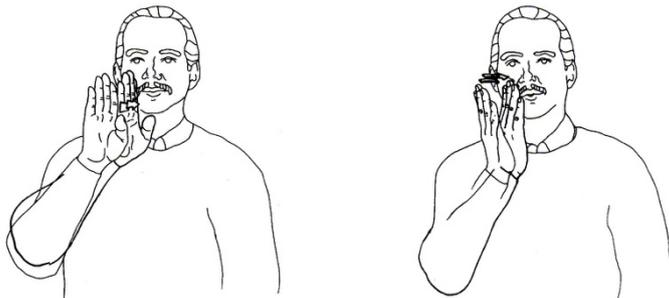
**Figure 8.** (different configuration) “bird” (Radutzky 2001:510.2) “to speak” (Radutzky 2001:534.2)



**Figure 9.** (different location) “France” (Radutzky 2001: 337.2) “thirteen” (Radutzky 2001:351.2)



**Figure 10.** (different movement) “family” (Radutzky 2001:300.2) “full” (Radutzky 2001:114.2)



**Figure 11.** (different palm orientation) “dad” (Radutzky 2001:91.3) “man” (Radutzky 2001:91.2)

It is also important to mention that a sign can be realized with one or two hands. One-handed signs are realized by left-handed signers with the left hand and by right-handers with the right hand. Two-handed signs are realized with two hands. In this case, both or only one of them may perform movement.

Different sign languages also vary in the cheremes they have in their systems. Some parameters are used in all SLs and others are not. Furthermore, some distinctions may be relevant in a particular sign language, but not in another. For example, the handshape W is very common in ASL and BSL, but not in LIS. Moreover, the handshapes in Figure 12 and Figure 13 are two different handshapes in LIS (A and S respectively), but not in ASL, where Figure 13 is a variant of Figure 12:



**Figure 12.** Handshape “A”  
(Caselli *et al.* 2006:68)



**Figure 13.** Handshape “S”  
(Caselli *et al.* 2006:68)

Additionally, a handshape may have different connotations in different sign languages. For instance, the handshape “I” has a negative connotation in BSL, but not in LIS (Caselli *et al.* 2006)<sup>5</sup>:



**Figure 14.** (BSL) Handshape “I”  
(Caselli *et al.* 2006:69)



“worse”



“wrong”

<sup>5</sup> The reader is referred to Boyes-Braem (1981) and Volterra (1987) for further research.

Thanks to Stokoe's analysis, we can claim that like spoken languages sign languages have a sub-lexical structure that is systematically organized. However, there is an important difference between phonemes and cheremes. The former are indeed sequential units, the latter are simultaneous. This is also due to the fact that spoken languages use the vocal-auditory modality to convey meaning, whereas sign languages operate within the visual-manual modality.

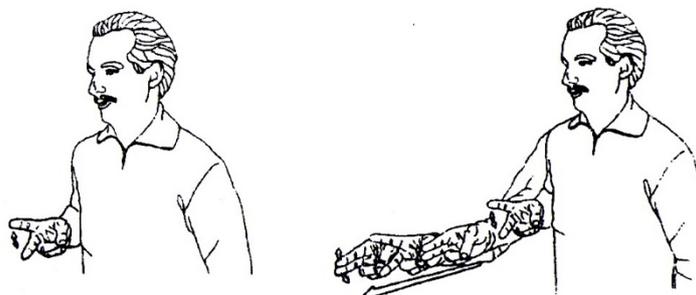
## **1.4 Morphology**

Morphology is the study of the internal structure of words. It deals with how the smallest units of meaning, that are morphemes, are combined to form words from elements like roots and affixes. In this context it is necessary to introduce a distinction between derivational (or lexical) and inflectional morphology (Matthews 1974). Specifically, derivational morphology concerns the relations among the linguistic elements (e.g. nouns and verbs), whereas inflectional morphology involves modifications in terms of gender and number. If we superficially compared spoken and sign language, we could easily claim that SLs lack a morphological system, since there seem to be neither inflection nor morphological markers that distinguish nouns from verbs and that allow creating new words from the pre-existing ones. In this paragraph I am going to show that this assumption is not true.

### **1.4.1 Nouns and verbs**

Supalla and Newport (1978) for ASL and Pizzuto and Corazza (1996) for LIS observe that nouns and verbs are not always semantically and morphologically related, but if they are, they refer to concrete objects and actions and can be divided into two subclasses. In this context, signs taken into consideration have to be analyzed in their citation form, as they were realized out of the blue and therefore not morphologically inflected.

As far as LIS is concerned, a subclass of signs includes noun-verb pairs that “share most but not all formational characteristics (e.g. handshape, orientation, place of articulation)” (Pizzuto and Corazza 1996:172) and the difference between them lies in the sign movement only: short and stationary for the noun, longer and with a specific direction for the verb, as we can see in Figure 15.



**Figure 15.** “scissors” “to cut with scissors” (Pizzuto 1987:185)

The second subclass comprises semantically related nouns and verbs, which have the same handshape, location, palm orientation and movement and can only be disambiguated through contextual information. Pizzuto and Corazza (1996) give for LIS the following example:



**Figure 16.** “telephone” or “to telephone” (Pizzuto and Corazza 1996:174)

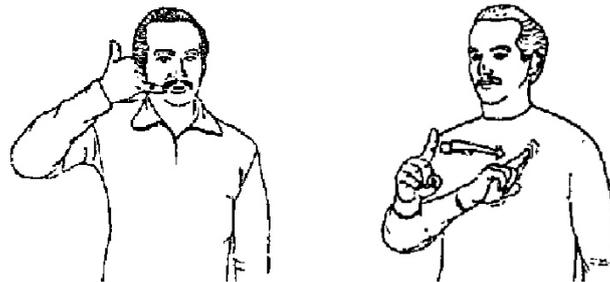
The sign represented in Figure 16 means either “telephone” or “to telephone”, depending on the context.

However, we can identify the distinction between the two different meanings. The sign is considered a verb if it is realized with a slow and repeated movement that confers the meaning of aspectual inflection, as can be seen in Figure 17 on the next page.



**Figure 17.** “to telephone for a long time” (Pizzuto and Corazza 1996:174)

The sign is considered as a noun if it is accompanied by a modifier such as “new” or by a possessive pronoun such as “my”, as shown in Figure 18:



**Figure 18.** “telephone” “my” = “my telephone”  
(Pizzuto and Corazza 1996:174)

#### 1.4.1.1 Noun morphology

Nouns in LIS can be divided into two morphological classes (Pizzuto 1987; Pizzuto *et al.* 1990; Pizzuto and Corazza 1996; Caselli *et al.* 2006).

The first class includes nouns that are articulated on the signer’s body, such as the sign for “woman” in Figure 19:



**Figure 19.** “woman” (Pizzuto 1987:186)

This kind of elements is always realized in their citation form. The only way to pluralize them is to add a sign that conveys the idea of plurality, such as the adverb “many”:



Figure 20. “woman”



“many” (Pizzuto 1987:187)

The second class comprises nouns that are articulated in neutral space, such as the sign for “town” in Figure 21. Their plural form is realized by displacing them in at least three points of neutral space, as in Figure 22:

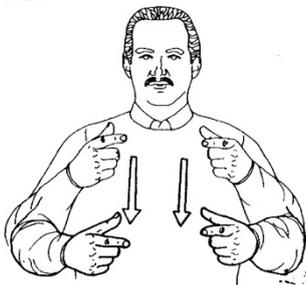


Figure 21. “town” (Pizzuto 1987:186)

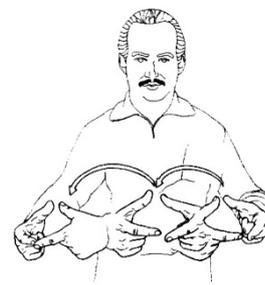


Figure 22. “towns” (Pizzuto 1987:187)

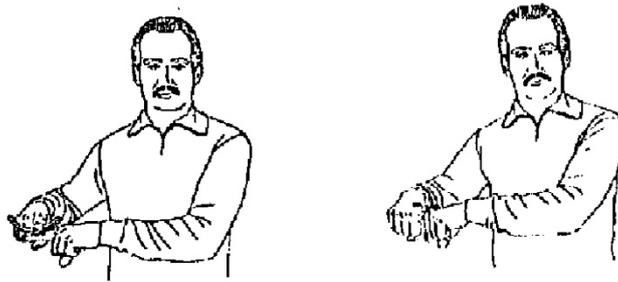
Furthermore, the modification of the place of articulation makes it possible for them “to specify or mark a nominal argument for deictic-anaphoric reference and for agreement of the verb with the same argument” (Pizzuto and Corazza 1996:176). Hence, the place in which they are articulated can become marked. Let us exemplify what this means with another sign of this class, such as “knife”:



Figure 23. “knife”  
(Pizzuto and Corazza 1996:177)



Figure 24. “to break”  
(Pizzuto and Corazza 1996:177)



**Figure 25.** “The knife is broken”  
(Pizzuto and Corazza 1996:177)

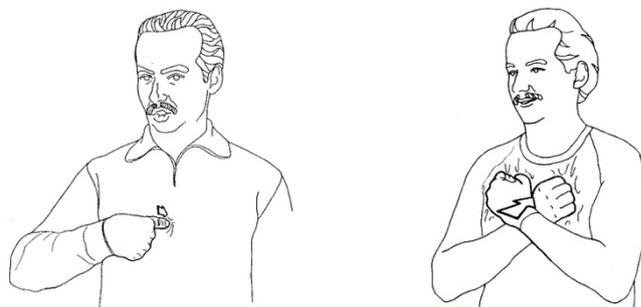
As we can see in Figure 25, the sign for “knife” is articulated in a marked position. Consequently, the verb “break” can be realized in the same position to express agreement with his argument.

This is not possible for nouns of the first class.

#### 1.4.1.2 Verbal morphology

According to Pizzuto (1986), (1987), Pizzuto *et al.* (1990) and Caselli *et al.* (2006), verbs in LIS can be divided into three classes.

The first class comprises verbs that are articulated on the signer’s body and do not allow modifications of palm orientation and movement direction, such as “believe” (Bertone 2011). Their citation form remains always invariable and they cannot agree with their arguments, as we can see in Figures 26 and 27.



**Figure 26.**

“I”

(Radutzky 2001:212.3)

“to believe”

(Radutzky 2001:653.1)

= “I believe”



**Figure 27.**

“He”/”She”

(Radutzky 2001:231.1)

“to believe”

= “He/She believes”

(Radutzky 2001:653.1)

Verbs of the second class are realized in neutral space. These signs express agreement with their arguments by varying their movement direction from one referential point to another:



**Figure 28.** “to take advantage of”

(Pizzuto 1987:190)



**Figure 29.** “He takes advantage of me.”

(Pizzuto 1987:193)

The third class includes verbs that are realized in neutral space, but mark only one referential point in it. Therefore they show overt agreement with one argument only. The argument is internal in unaccusative and transitive verbs and external in the intransitive ones (Pizzuto 1986). In this case, the direction of the movement does not vary, whereas the place of articulation does:



**Figure 30.** “to grow up”

(Pizzuto 1987:195)



**Figure 31.** “A/The child grows up.”

(Pizzuto 1987:195)

From Figure 31, we can easily deduce that the movement direction establishes syntactic relations between the discourse referents.

### 1.4.2 The use of space

Unlike spoken languages, which use a vocal modality, sign languages convey meaning through the visual channel. For this reason, it is no wonder that many signs are produced in the area in front of and at the side of signers, which is the signing space. More precisely, the signing space extends from the top of the head to the waist, and from shoulder to shoulder.

So far, researchers have considered it neutral from a phonological point of view, since there are no minimal pairs that differ in the place of articulation in the space (Verdirosi 1987). However, the same statement does not apply for the role of space in morphology and syntax. In other words, the realization of the sign in a particular point in the space can have a specific purpose. 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 (as discussed in Section 1.4.1). The latter is related to specification marking: the ideas of specificity and definiteness imply the identification of a particular point in the space that differs from the others.

Taking this into account, we can distinguish a referential, defined space from a neutral, undefined space (Bertone 2009)<sup>6</sup>.

As for LIS, Bertone (2009) considers points of space as grammatical *loci* that refer to the verb arguments. Consequently, we can claim that location as formational parameter can define thematic relations of verbs:

(1) <sub>1</sub>TAKE-ADVANTAGE-OF<sub>2</sub> <sup>7</sup>

“I take advantage of you.”

In (1), the places of initial and final articulation are indicated with <sub>1</sub> (first person singular) and <sub>2</sub> (second person singular) and mark agent and patient, respectively.

---

<sup>6</sup> See Klima and Bellugi (1979), Padden (1990), Bahan (1996), Meier (1990), Liddel (1995, 2002) for further research.

<sup>7</sup> I adopt the common notational convention to gloss manual signs of Sign Language data in capitals. It is also possible to gloss them in small capitals. Non-Manual Markers are eventually annotated in the line above the manual sign gloss. For further information the reader is referred to Section 1.8.4.

If there are overtly realized arguments, they have to be co-indexed with the places of articulation (2). If they are not, the sentence is ungrammatical (3).

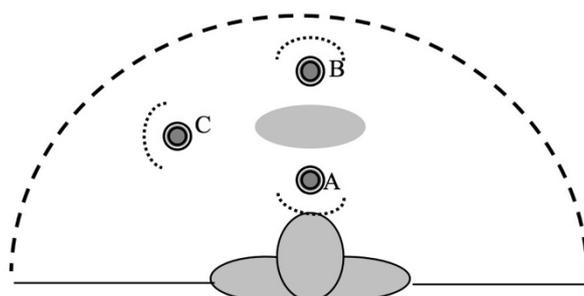
(2) IX-1<sub>1</sub> JOHN<sub>3</sub> 1TAKE-ADVANTAGE-OF<sub>3</sub>

“I take advantage of John.”

(3) \* IX-1<sub>1</sub> JOHN<sub>3</sub> 2TAKE-ADVANTAGE-OF<sub>3</sub>

Let us now focus on the representation of the different places of articulation, which remains a problematic and widely discussed topic. First of all, I would like to stress that points in the space do not usually preserve the same reference and do not usually convey the same meaning in different contexts. 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 32 (Mac Laughlin 1997) shows how the representation works. On this basis, we can outline some space features.



**Figure 32.** Top view of the signing space

Points in space have two semantic features: proximal [ $\pm$ prox] and distal [ $\pm$ 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 and definite. It is specific because it is bound to a referent; it is definite because it can be identified. Consequently, an unspecified point in space is undefined, therefore [-dist] (Bertone 2009).

In the following table the reader can find my translation of the table showed in Bertone (2009:85) that sums up semantic features and their meaning:

SPATIAL FEATURES	PLACE OF ARTICULATION FROM SIGNER'S POV <sup>8</sup>	MEANING
[+prox]	Defined space in front of the signer	here, this, now, I*, my*
[-prox]	Defined space between signer and interlocutor	there, near you, tomorrow, future, you*, your*
[+dist]	Defined space far from the signer and the interlocutor, also marked by gaze direction that does not coincide with the direction from the signer to the interlocutor	there-that, historical time (one day, etc.), he/she*, his/her**
[-dist]	Undefined space	where? anywhere, everywhere, always-never, someone

\* personal features specified by pointing toward the signer have to be added to the spatial features.

\*\* possession is specified by handshape and palm orientation that differ from the other signs.

**Table 1.** Semantic features and their meaning

As previously mentioned, I usually annotate sign language data manually. I also utilized this technique to transcribe the fable “The Tortoise and the Hare” from LIS to Italian for the present work. In this narrative, there are many different locations and pointing signs and I decided to gloss them considering their semantic features (please see Chapter 3).

According to Cormier *et al.* (2013), pointing with an extended index finger can be used with many purposes within sign languages. For instance, the act of pointing could be interpreted as locative adverb, pronoun or determiner, depending on the context. Moreover, it makes spatial features explicit (Bertone 2009).

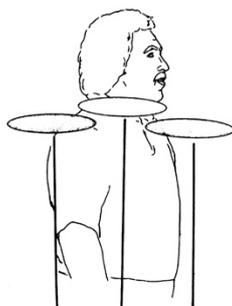
---

<sup>8</sup> POV stands for Point Of View.

### 1.4.3 Time, tense and aspect

Many authors<sup>9</sup> have proven that sign language have tenses.<sup>10</sup>

In general, signers use an imaginary time line as a time indicator to express past, future and present tenses: past events are located in the space behind the signer, future events in the space in front of the signer and the space in which the signer is located represents the present. Therefore, we can claim that the signer's shoulder counts as temporal boundary line:



**Figure 33.** past present future

According to Zucchi (2009), temporal information can be conveyed in LIS through:

- i. Adverbs of time that in LIS shift the speech point (s-point)<sup>11</sup>, that is the tense of the verb, as in:

(4) YESTERDAY MARIA ICE-CREAM EAT

“Yesterday Mary ate an/the ice-cream”

(5) TOMORROW GIANNI MARKET GO

“Tomorrow John will go to the market”

In this kind of LIS sentences, tense is present and equates the e-point<sup>12</sup> with the s-point shifted by the adverb.

- ii. NMMs co-occurring with the verb, such as a particular position of the signer's shoulder:

---

<sup>9</sup> E.g.: Jacobowitz and Stokoe (1988), Neidle, Kegl, Mac Laughlin, Bahahn, Lee (2000) for ASL; Zucchi (2009) for LIS.

<sup>10</sup> There are contrasting opinions on this topic. Some authors suppose that sign languages do not have grammatical tense. See Friedman (1975) for ASL and Pizzuto *et al.* (1995) for LIS.

<sup>11</sup> Zucchi distinguishes between s-point and time of utterance. The former equates the grammatical tense, whereas the latter corresponds to the moment in which the sentence is uttered.

<sup>12</sup> The e-point is the time of the event described by the sentence. For further information on the parameters that order relations in time the reader is referred to Reichenbach (1947).

shoulder straight

(6) GIANNI HOUSE BUY  
 “Gianni is buying a house”

shoulder backward

(7) GIANNI HOUSE BUY  
 “Gianni bought a house”

shoulder forward

(8) GIANNI HOUSE BUY  
 “Gianni will buy a house”<sup>13</sup>

As we can see in (6-8), if the shoulder is aligned with the body during the realization of the verb, the verb is considered to be in the present tense. If the shoulder is tilted backward, the verb is in the past form and if the shoulder is tilted forward, the verb is in the future form. In other words, shoulder’s position can be a way of inflecting the verb for tense.

The author also demonstrates that past and future tenses in LIS are absolute tenses and they cannot co-occur with time adverbs, as shown in (9) and (10):

shoulder backward

(9) \* TIME-AGO GIANNI HOUSE BUY (Zucchi 2009:114)

shoulder forward

(10) \* TOMORROW GIANNI HOUSE BUY (Zucchi 2009:114)

iii. lexical markers such as DONE (11) that indicates that the action expressed by the verb is completed, and MUST (12) that indicates that the action will take place in the future:

(11) MARIA EAT DONE  
 “Maria has eaten”

(12) MARIA BOOK BUY MUST  
 “Maria must buy a/the book”

---

<sup>13</sup> Examples (6-8) are showed in Zucchi (2009:101).

iv. the context (13):

(13) YESTERDAY GIANNI MOVIE-THEATER GO THERE MARIA HIM  
MEET

“Yesterday Gianni went to the movie-theater. Maria met him there”

(Zucchi 2009:102)

As Zucchi (2009) points out, the action of the first sentence in (13) occurred in the past. Consequently, the tense of the second sentence is interpreted as past, although it does not display an overt marker of past time.

Bertone (2011) provides an overview of the aspect in LIS and distinguishes between perfective and imperfective aspect, as in Bertinetto (1986, 1991).

Perfective aspect in LIS describes a process viewed as a whole and as a concluded event. It can be expressed through a verb followed by the sign DONE.

Imperfective aspect refers to ongoing or not concluded processes. It can be expressed through variation of the duration of the sign, repetition of the sign and with adverbs of time and frequency.

#### 1.4.4 Classifiers

In this section we focus on one of the most fascinating peculiarities of sign languages: the use of classifiers.

Classifiers (CLs) are grammatical units used in many spoken languages that classify nouns according to specific semantic classes<sup>14</sup>. 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 languages<sup>15</sup> and do not critically differ in their functions from classifiers in spoken languages. Classifiers in SLs are usually to be found in post-nominal position (Mazzoni 2012) and are defined as

---

<sup>14</sup> For further reading on the classifier systems in spoken languages, the reader is referred to Worsley (1954); Dixon (1968, 1982); Denny (1976); Allan (1977); Craig (1986, 1992) and Aikhenvald (2000).

<sup>15</sup> The reader is referred to Schembri (2003:28 – endnote 1) and Mazzoni (2008:35) for the above-mentioned studies.

specific handshapes realized in combination with specific predicative roots, i.e. movements in conjunction with handshapes. From this, we can easily infer that they are not standard, but rather variable structures. This also accounts for the fact that they are usually not reported in sign language dictionaries.

Caselli *et al.* (2006) observe that LIS signers often create new signs making extensive use of classifiers in order to convey meaning. CLs are also utilized in narratives and poems, since they allow expressing images and figurative concepts that are difficult to translate with words (Corazza 1990, Pizzuto and Corazza 1996). According to Supalla (1986), they are among the most iconic elements of sign languages.

CLs have been first identified by Frishberg (1975) in a study on ASL and further investigated by various researchers, who suggested different criteria for their classification.

As far as LIS is concerned, a first analysis of classifiers belongs to Corazza (1990), who follows the model of Liddell and Johnson (1987). Let us focus instead on other two analysis: Bertone (2011) and (Mazzoni 2012).

Bertone (2011) adopts the categorization proposed by Aikhenvald (2000) for spoken languages and distinguishes:

- i. Noun classifiers, which describe the form of a noun and are selected on the basis of their semantic features. Figure 34 shows the classifiers of a plane leaf, a pine needle and an olive leaf, respectively. They clearly have iconic components:



**Figure 34.** PLANE-LEAF PINE-NEEDLE OLIVE-LEAF  
(Bertone 2011:63)

- ii. Locative or deictic classifiers, which establish spatial relations among different elements. They are deictic when they define a specific referent in relation to

others. For example, Figure 35 shows the classifiers of two referents positioned next to each other, which in context of Figure 35 are a hare and a tortoise:



**Figure 35.** Classifiers for HARE and TORTOISE<sup>16</sup>

iii. Numeral classifiers, which are used to inflect nouns. If used in combination with plural nouns, they convey information on their number and position in space. In Figure 36 below, the signer uses a classifier with the handshape “4”, which denotes plurality of long referents. In this case, referents are people:



**Figure 36.** CL: 4 people standing in line CL: 4 people standing in line (Bertone 2011:66)

iv. Verbal classifiers, which are used as verbs. In fact, the lexical root, i.e. the handshape, incorporates the predicative morpheme, i.e. the movement.



**Figure 37.** moving boat (spreadthesign.com)

<sup>16</sup> Figures with this signer in the present thesis 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.

Mazzoni (2012) suggests a new model based on the studies of Enberg-Pedersen (1993) on Danish Sign Language and Benedicto and Brentari (2004) on ASL. The author classifies classifiers into four semantic categories:

- i. Whole Entity Classifiers, which include handshapes that represent the nominal referent as a whole:



**Figure 38.** Classifier for “toothbrush” (Mazzoni 2012:49)

- ii. Handling Instrument Classifiers, which indicates how the hand holds an object while holding and manipulating it:



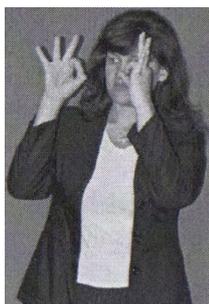
**Figure 39.** Classifier for “holding a toothbrush” (Mazzoni 2012:51)

- iii. Extent and Surface Classifiers, which refer to a specific characteristic or feature of the nominal referent:



**Figure 40.** Extension classifier for sphere shaped objects (Mazzoni 2012:52)

- iv. Limb/Body Part Classifiers, which denote specific parts of a body, such as head, legs, feet, eyes, tongue, but also paws, tail, crest and horns.



**Figure 41.** Classifier for “eyes” (Mazzoni 2012:53)

She observes that the handshapes are realized in combination with the following predicate 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, like in Figure 42:



**Figure 42.** Classifier for “person moving” (Mazzoni 2012:55)

- ii. Manner or imitation roots, which represent an action or the type of movement of the referent:



**Figure 43.** Classifier for “to zigzag” (Mazzoni 2012:56)

- iii. Position or contact roots, in which the combination movement-hold produced by the hand situates the referent in space:



**Figure 44.** Classifier to indicate a vehicle's position (Mazzoni 2012:56)

- iv. Extension or stative-descriptive roots, in which the hand movement shows the state of the referent or illustrates how referents are positioned in space:



**Figure 45.** Classifier for "homogeneous elements/referents in a circle" (Mazzoni 2012:57)

On the whole, we can claim that these studies have an essential role, since they prove that sign language have complex linguistic structures. They also demonstrate that iconic elements, such as classifiers, can display linguistic meaning.

## 1.5 Syntax

Syntax studies structure and classification of phrases, clauses and sentences, including word order. There is not just one word order, but an order is usually more basic than others and has some specific characteristics. According to Brennan (1994) and Dryer (2007), the more basic order is the most frequent order, used in simple declarative active clauses. Furthermore, it is the least morphologically and pragmatically marked.

Constituent word orders are defined in terms of a finite verb (V) in combination with the subject (S) and the object (O). The most common orders for spoken languages are SOV, SVO and VSO. VOS, OSV and OVS are more rare (Greenberg 1966; Hawkins 1983; Dryer 1989). Starting from the early 1970s, researchers began to study word order in sign languages and to identify the same variety of orders observed in spoken languages (Branchini and Geraci 2010).

### 1.5.1 Word order

As far as LIS is concerned, the first pioneering studies belong to Volterra *et al.* (1984), Laudanna (1987) and Laudanna and Volterra (1991).

Laudanna (1987) analyses two sets of data, acquired from an elicited production task and a grammatical judgement task, respectively.

The former displays that the order is usually SVO in reversible sentences and mostly SOV in non-reversible sentences.<sup>17</sup> SOV order is also preferred when the signer realizes classifiers or uses the space to mark agreement between the arguments.

The latter shows that SVO is the most acceptable linear order, OSV and VOS are marked orders. SOV is also accepted, if the subject is followed by a pause or if the verb incorporates the object.

The linear order of LIS has been further investigated by Geraci (2002) and Branchini and Geraci (2010).

Contrary to what Laudanna (1987) observes, Geraci (2002) demonstrates that SOV is the basic linear order in LIS. SVO is acceptable in sentences that contain complex objects, like in (14):

(14) DANIELE SERGIO GIFT BOOK BOUGHT SHOP CITY-CENTRE  
SCHOOL

NEAR

“Daniele gifts to Sergio the book bought in the city centre near the school”

(Branchini and Geraci 2010:118)

---

<sup>17</sup> If subject and object are reversed and the sentence remains meaningful, the sentence is reversible (“The dog chases the cat” – “The cat chases the dog”). By contrast, in a non-reversible sentence arguments cannot be swapped (“The mother ties Anna’s laces” - \* “Anna’s laces tie the mother”).

Lastly, the research conducted by Branchini and Geraci (2010) reveals that the SOV order is slightly more used than SVO (54% vs. 46%). However, the use of these orders is conditioned by linguistic factors, such as the presence of functional signs and the reversibility of the predicate; and by social factors, such as the geographical origin of the signers.

### 1.5.2 Other general properties

The grammatical judgement task proposed by Laudanna (1987) shows that temporal markers like TOMORROW are located at the beginning of the sentence:

- (15) TOMORROW I CINEMA GO  
 “Tomorrow I’ll go to the cinema” (Laudanna 1987:217-218)

However, adverbs like ON-TIME can be found at the end of the sentence:

- (16) GIANNI ARRIVE ON-TIME  
 “Gianni arrived on time” (Geraci 2006:218)

In negative sentences, negation is located at the end of the sentence:

- (17) I HOME GO NOT  
 “I don’t go home” (Laudanna 1987:218)

This has also been confirmed by Geraci (2006), who provides a detailed study of negation in LIS, in which he identifies the negative markers NON (as “not”) and NEG (as “presuppositional not”) and the negative words NOBODY and NOTHING. The author remarks that these negative elements have to be accompanied by specific negative Non-Manual Markers (18), otherwise the sentence is ungrammatical.

- (18) PAOLO CONTRACT SIGN <sup>neg</sup> NON  
 “Paolo didn’t sign the contract” (Geraci 2006:221)

They generally occur post-verbally; therefore they follow verb and modal verb. However, it is possible to find a negative marker in preverbal position. In this case, NMMs spread rightward over the rest of the sentence, as shown in (19):

- (19)  $\frac{\text{neg}}{\text{NOBODY CONTRACT SIGN}}$   
 “Nobody signed the contract” (Geraci 2006:221)

The aspectual marker DONE and modal verbs follow the lexical verb:

- (20) GIANNI HOUSE BUY DONE  
 “Gianni bought a house” (Geraci *et al.* 2008:47)

- (21) GIANNI METER 80 JUMP CAN  
 “Gianni can jump 1.80 mt.” (Geraci *et al.* 2008:47)

WH-phrases are found in the right periphery of the sentence and are characterized by specific Non-Manual Markers, that spread over the WH-phrase:

- (22)  $\frac{\text{wh}}{\text{GIANNI BUY WHAT}}$   
 “What did Gianni buy?” (Cecchetto *et al.* 2009:282)

- (23)  $\frac{\text{wh}}{\text{HOUSE BUY WHO}}$   
 “Who bought a house?” (Cecchetto *et al.* 2009:282)

If the WH-phrase is subject, they may optionally spread over the whole sentence (24). If it is object, they spread over the whole sentence except the subject (25) (Cecchetto *et al.* 2009):

- (24)  $\frac{\text{wh}}{\text{CAT CHASE WHO}}$   
 “Who chases the cat?”

- (25)  $\frac{\text{wh}}{\text{CAT CHASE WHOM}}$   
 “Whom chases the cat?”

The reader is referred to Bertone (2011) and Branchini (2014) for further general reading on LIS. Please see Geraci (2006), Geraci *et al.* (2008), Cecchetto *et al.* (2009) for studies on negation, sentential complementation and right periphery, respectively.

## 1.6 Non-Manual Markers<sup>18</sup>

People say that a picture is worth a thousand words; I would say that a facial expression is worth a thousand words too, or even more. Let us focus on one of the most underestimated components of sign languages: Non-Manual Markers.

SLs are expressed through manual articulations that are realized in combination with NMMs. Non-Manual Markers include facial expressions, movements of the body, head, eyebrows, eyes, cheeks, mouth, shoulders and postures and contribute to convey information together with manual signs.

It is possible to identify different functions of NMMs. First of all, we need to distinguish between non-linguistic and linguistic NMMs, even if it is not always easy. In broad terms, non-linguistic NMMs are affective and are used to convey emotions, whereas linguistic NMMs have specific grammatical functions. 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). Their realization is obligatory.

On the contrary, affective expressions are facultative, 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, 1980; Emmorey 1999; Wilbur 2000; Herrmann 2013). They could also be considered part of the pragmatic level in sign languages.

The difference between these two types of Nonmanuals has also been proved from a neurological point of view (Corina *et al.* 1999).

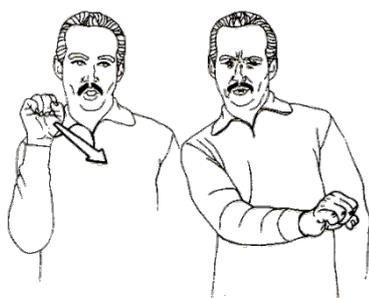
Since NMMs convey both affective and linguistic information using the visual channel only, they have been compared to intonation in spoken languages (Reilly *et al.* 1992; Nespor and Sandler 1999).

---

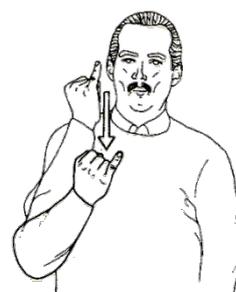
<sup>18</sup> Also known as Non-Manual Features (NMFs), Non-Manual Components (NMCs) or just as Nonmanuals.

Let us now focus on linguistic NMMs, which can be distinguished in lexical, adverbial and syntactic.

If we analyse them from a phonological point of view, we can notice that they necessarily occur in simultaneous conjunction with some manual signs. For instance, the signs for emotions, such as HATE (Figure 46) and physical characteristics, such as THIN (Figure 47) have to be produced in LIS with a particular facial expression that completes the meaning of the manual sign. We refer to this kind of components as lexical NMMs.

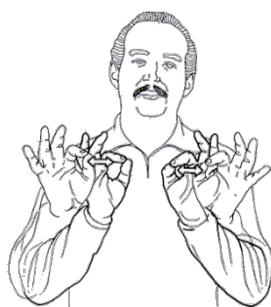


**Figure 46.** “hate” (Franchi 1987:160)

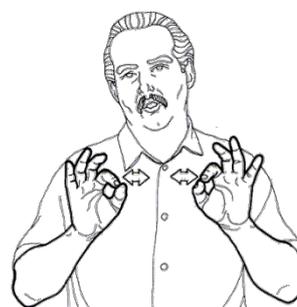


**Figure 47.** “thin” (Franchi 1987:160)

I have already mentioned that NMMs are considered the fifth formational parameter in sign languages. In LIS, lexical Nonmanuals can have indeed a distinctive function. For example, the signs for “fresh” and “not yet” are both realized with the same manual features (hand configuration, location, orientation and movement). Nonetheless, they convey two different meanings that can be distinguished only through NMMs:



**Figure 48.** “fresh” (Radutzky 2001:693.3)



**Figure 49.** “not yet” (Radutzky 2001:694.1)

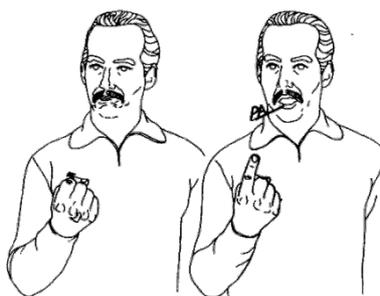
Two pioneering studies on NMMs involving the mouth (Vogt-Svendsen 1984; Schroeder 1985) identified two types of oral components that are coextensive with

manual signs: Borrowed Word-Pictures and Special Oral Components<sup>19</sup>. As far as LIS is concerned, Franchi (1987) distinguishes analogously between *Immagini di Parole Prestate* (IPP) and *Componenti Orali Speciali* (COS). IPP include mouth patterns that silently articulate an emphasized segment of the corresponding word, whereas COS produce mouth gestures that are connected to proprioception and are completely independent from spoken languages. Figure 50 is an example of a sign with IPP, in which the signer produces the sign AVVOCATO, which means “lawyer”, and articulates the salient letter V with the mouth.



**Figure 50.** “lawyer” (Franchi, 1987:163)

Figure 51 shows a COS: during the realization of the sign, the signer inflates his cheeks and then releases the air from the mouth.



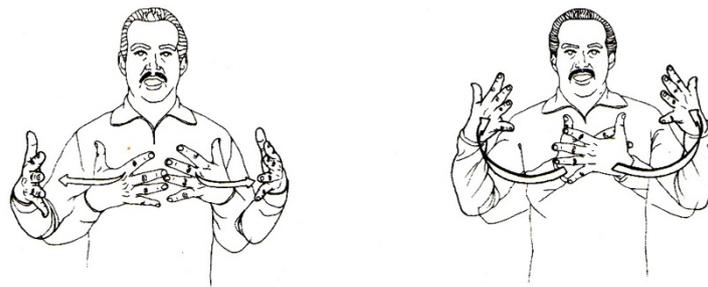
**Figure 51.** “come out” (Franchi 1987:162)

NMMs play a crucial role within the morphological domain too, since they can inflect the verb for tense (Section 1.4.3) and modify adjectival and adverbial information.

For instance, the sign VERY-BIG in LIS is a modification of the adjective BIG: the hands of the signer realize a wider movement and his eyes and mouth are opened. As we can see in Figure 52, this is not the case for BIG.

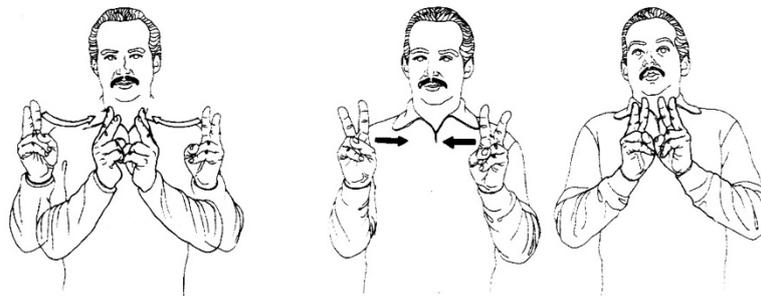
---

<sup>19</sup> The reader is referred to Boyes Braem and Sutton-Spence (2001) for further study.



**Figure 52.** “big” “very big” (Franchi 1987:165)

Another example is provided in Figure 53, that shows the realization of the verb TO-MEET and TO-MEET-SUDDENLY. As can be seen, the sign is articulated with a stretched and fast movement, accompanied by a particular expression that is consistent with the adverbial meaning.



**Figure 53.** “to meet” “to meet suddenly” (Franchi 1987:162)

Moreover, nonmanuals have another morphological function, since they are involved in verb agreement (Bahan 1996). In fact, the head of the signer is usually tilted towards the position in the signing space associated with the subject and the gaze is directed towards the position associated with the object, as can be seen in the following LIS sentence:

$\overline{\text{HEAD}_k}$   
 $\text{GAZE}_i$   
 (26)  $\text{GIANNI}_k \text{ MARIA}_i \text{ LOVE}$   
 “Gianni loves Maria”

Last but not least, non-manual markers are also essential within the syntactic domain, since they determine the sentence type, mark topicalized constituents and are 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 allow the realization of conditional and relative constructions. Let us briefly focus on them one by one.

The only difference between a declarative and a yes/no question lies in NMMs. Indeed, polar questions are realized with raised eyebrows, wrinkled forehead and head and shoulders tilted forward (Franchi 1987).

WH-questions are characterised by a WH-element in the right periphery and furrowed eyebrows. These specific NMMs obligatory spread over the WH-element. However, they may spread in a wider way: all over the sentence if the WH-element is subject or all over the sentence except the subject if the WH-element is object (Cecchetto *et al.* 2009).

Non-manual features in LIS negative sentences consist of lowered eyebrows and a headshake obligatory co-occurring with the negative manual sign. They usually do not spread over broader domains (Geraci 2006).

Imperative sentences are characterized by specific imperative manual and non-markers. NMMs obligatory spread over the whole sentences and differ on the basis of the type of imperative meaning (commands, suggestions, invitations, etc.) they convey (Donati *et al.* 2017).

Topic and focus phrases have in LIS specific nonmanuals, too. Raised eyebrows and head nod for the former, followed by eye-blinking; raised eyebrows, tense and wide open eyes and lifted chin for the latter.

NMMs in conditional constructions are raised eyebrows, wide open eyes and forward tilted head and shoulders (Franchi 1987), and spread over the entire dependent clause.

Relative constructions are sentences characterised by the presence of the sign PE, which is realised at the end of the subordinate sentence and refer to a constituent in it. Typical NMMs are raised eyebrows and tense eyes and upper cheeks and extend over the relative clause (Cecchetto *et al.* 2009; Branchini 2007; Branchini and Donati 2009).

So far we have seen that Non-Manual Markers obligatory have to co-occur with manual signs. What if it is not always so?

Dively (2001) investigated in ASL the so-called “nonhanded signs” (NHSs), namely signs realized without using the hands. The author identifies eight NHSs 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. The eight mentioned NHSs are:

- i. NHS:YES, meaning “yes”;
- ii. NHS:NO, for “no”;
- iii. NHS:THEN, meaning either “ready” or “next in a sequence of events”, depending on context;
- iv. NHS:OH-I-SEE, for “oh, I see” or “yes, I do understand”;
- v. NHS:WRONG, which indicates that the previously conveyed information was not correct;
- vi. NHS:OR, for “or”;
- vii. NHS:PUZZLED, meaning either “I am puzzled” or “that’s strange”;
- viii. NHS:TIME-PASSED-BY, meaning “time passed by” (Dively 2001).

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.

In this regard, there is another point that I would like to stress. I think that there are contexts in which signers convey meaning without using the hands and it is not about discourse structural components. This can happen through Role Shift, which also allows the signer to express (and report) simultaneous and different information with both manual and non-manual articulators. Let us get to the heart of the matter in the next Section.

## 1.7 Role Shift

I would like to write a few lines on one of the most productive constructions in SL: Role Shift (RS).

Role Shift<sup>20</sup> 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 communities<sup>21</sup>, but not very common in spoken languages. It is characterised by specific body markers, which according to Mazzoni (2012) are:

- i. Precise position of all the referents within the space, which becomes a real narrative setting;
- ii. Temporary interruption of eye contact with the actual interlocutor 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).

It is also important to notice that RS is not introduced by particular expressions. For this reason, we have to consider it as a grammatical phenomenon that differs from direct speech in spoken languages (Ajello 1997; Quer 2016).

Role Shift can be used to report a speech or thought of a referent or to reproduce his or her actions, thus it can be divided into two varieties (Schlenker 2017a; Schlenker 2017b). The terminology for both phenomena is not consistent throughout the

---

<sup>20</sup> Also known as “referential shift” (Emmorey and Reilly 1998), “constructed action” (Metzger 1995), “context shifting operator” (Zucchi 2004).

<sup>21</sup> E.g. it 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 (1993; 1995); in South African Sign Language by Aroons 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 *et al.* (2005) and Mazzoni (2008; 2009; 2012).

literature<sup>22</sup>. I adopt in this work the terminology used by Herrmann and Pendzich (2018), namely “quotation role shift” (Q-RS) and “action role shift” (A-RS).

Hence, Q-RS is the type of RS by which the signer reports words or thoughts of other referents. A-RS allows the iconic reproduction of actions, mannerisms and emotional states, including facial expressions and non-linguistic gestures. It should be clear that it is not a description, but rather a demonstration (cf. Clark and Gerrig 1990:764-768). It is not even pantomime, since it is conveyed only with upper parts of the body (e.g. torso, head, eye gaze). Furthermore, in A-RS a large variety of classifiers is often used. For instance, when embodying an animal that stands still, the signer may manually realize classifiers that represent the limbs of that animal, as Figure 54 shows:



**Figure 54.** Role Shift representing a tortoise that stands still (LIS)

However, there is some overlap between both uses, since signers may use affective facial expressions of the referent whose utterance they report and use quotation during a phase of A-RS (Metzger 1995; Pfau and Quer 2010; Herrmann and Pendzich 2018).

As already mentioned, Role Shift is a widely used strategy in the Deaf community, especially in narratives. Nevertheless, the type of conveyed information is not binding. In other words, it is not a trigger for RS. For instance, signers do not tell stories using only A-RS: the use of A-RS depends on the chosen narrative strategy (Herrmann and Pendzich 2018), on the interlocutor and on the level of formality of the setting (Earis and Cormier 2013).

---

<sup>22</sup> For 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, 2017b) distinguishes between “attitude role shift” and “action role shift”.

Moreover, Q-RS and A-RS often co-exist and interact between each other in Sign Language narration. Signers use Role Shift and adopt the perspectives of narrator and protagonists. The perspective shift is realized rapidly and involves the simultaneous use of manual and non-manual articulators (Herrmann and Pendzich 2018).

At this point, it should be clear that RS is a very complex phenomenon. Let us consider for example Figure 55:



**Figure 55.** Example of A-RS (LIS)

Figure 55 shows an example of A-RS: the signer is embodying a haughty hare<sup>23</sup> that was hopping and looking around. This phase of A-RS gives us three simultaneous pieces of information. One: the hands articulate classifiers that represent the hare's limbs and their movement let us understand that the hare is hopping. Two: the right-to-left head movement is the head movement of the hare and shows indeed that the animal is looking around or even, to be more precise, that the gaze is focused first rightwards and then leftwards. Three: the specific affective facial expression represents the hare's haughtiness.

So, by seeing the head movement, we understand that the signer is conveying a particular meaning. I wonder if we should consider these structures as nonhanded signs, even if they probably convey the same meaning in different SLs and are not universal as Dively (2001) observed. Actually, this is exactly what I did in the annotation of the fable "The Tortoise and the Hare" for this thesis, to facilitate the correct and accurate generation from LIS to Italian.<sup>24</sup> Otherwise, a multilevel structure for A-RS that distinguishes among different conveyed information is needed.

Personally, I think that further research is necessary to clarify potential analogies between non-manual articulations during RS-phases and NHS.

---

<sup>23</sup> N.B.: the hare had already been introduced in context.

<sup>24</sup> I will discuss this topic in Chapter 3.

## 1.8 Writing systems for Sign Language

We live in a world where writing is everywhere. Writing is the primary basis upon which we learn, work and often communicate. It allows us to preserve our language. Until now, none of the existing sign languages seems to have autonomously developed a written form (Pizzuto *et al.* 2006). Thus, SLs continue to be passed down face to face from generation to generation.

Many attempts have been made to create a formally adopted written system, but most of them are not usable by the Deaf community, since they consist of complex sets of symbols and numbers.

In this section, I will introduce five main systems used for phonological or morphological notation of sign languages and the major annotation tools.

### Stokoe Notation

Stokoe Notation is a phonemic script for ASL developed by Stokoe (1960). It writes from the signer's point of view and it is based on a set of arbitrary symbols that notate the three parameters into which he decomposed signs. Stokoe assembled these aspects with specific terms: he named location *tabula* or *tab*, handshape *designator* or *dez* and movement *signation* or *sig*. These have to be written in the following order: *tab - dez - sig* (Frishberg *et al.* 2012), left to right. Figure 56 shows an example of the Stokoe notation for the ASL:

∩5'5"      Ø5 v" O o O o °□°∞ 55><  
grandfather child give                      ball  
'Grandfather gives (or gave) the child the ball.'

**Figure 56.** Example of Stokoe Notation in ASL (Frishberg *et al.* 2012:1049)

As we can see, the symbols used in this notation are Latin alphabet letters, numbers and some symbols invented ad hoc.

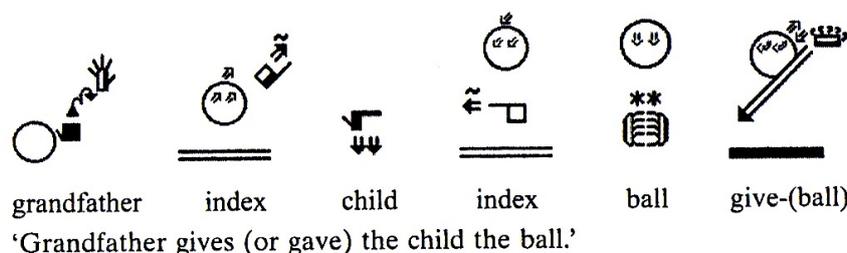
However, this analysis does not consider morphological processes, movement of the body in space and Non-Manual Markers. Furthermore, this system was meant for decontextualised signs and dictionary entries, and not for signs realized in conversation.



transcription of signed discourse and fail to accurately represent spatial locations and directions. They also provide a linear representation of constituents, which is only suitable for words and not for signs, since signs have multidimensional and multilineal features. Furthermore, they mainly describe manual components of signs and do not consider NMMs appropriately. They use also special fonts and cannot function without them (Hoiting and Slobin 2002; Pizzuto *et al.* 2010; Garcia and Sallandre 2013).

## Sign Writing

Sign Writing (SW) is “an alphabet for sign language”<sup>25</sup> invented by Sutton (1999), a “sign orthography”<sup>26</sup> that provides a set of graphic and highly iconic international symbols for hands, face, body with notations for location and direction. Hence, it allows providing a representation of co-occurring manual and non-manual components. It is written in vertical columns from the signer’s point of view (Frishberg *et al.* 2012). Figure 58 shows an example of SW in ASL (presented horizontally to save space):



**Figure 58.** Example of SW in ASL (Frishberg *et al.* 2012:1052)

SW was conceived with the idea of allowing people to write and read dictionaries, newspapers, books etc written with signs (Di Renzo *et al.* 2011).

This system has been adapted for many sign languages, including LIS, and seems to be suitable for the adoption as writing system by Deaf people:

“The deaf LIS signers who are using SW found, in the first place, that in spite of the apparent complexity of this writing system (it provides more than 35.000 potentially usable glyphs), they learned to use it very easily and rapidly.” (Pizzuto *et al.* 2010:226)

<sup>25</sup> Pizzuto *et al.* (2010:225)

<sup>26</sup> Crasborn (2015:76)

It is used in educational contexts and renowned in around 40 countries (Pizzuto *et al.* 2010; Garcia and Sallandre 2013).

However this system does not allow representation of morphological information and often provides multiple representations for the same sign (Frishberg *et al.* 2012; Garcia and Sallandre 2013). Furthermore, symbols used are not easily suitable for Natural Language Generation (NLG) and Machine Translation (MT).

### Manual glossing

The process of glossing consists in the transcription of information conveyed in Sign Language with words of a spoken language. It is the system I adopted to annotate the LIS fable “The Tortoise and the Hare” for this thesis.

Glosses are usually written in two lines: manual signs are transcribed in a line in capitals and eventually enriched with additional information about direction, movement and repetition of the sign. Nonmanuals that co-occur with the sign have to be transcribed too: their function has to be written on a line above the manual sign gloss, generally in small caps:

(27) <sup>topic</sup>  
BOOK , I READ DONE (LIS)  
“As for the book, I have read it”

(28) GIANNI<sub>j</sub> MARIA<sub>k</sub> BOOK<sub>j</sub>GIVE<sub>k</sub> (LIS)  
“Gianni gives Maria the book”

It is a subjective and fairly free transcription system: who is glossing can decide how to do it, what to transcribe and what to omit. Hence, there is no actual standardization. Furthermore, this kind of transcription often cannot be interpreted correctly without the provision of the corresponding pictorial material, as Frishberg *et al.* (2012) observe.

On the whole, we can claim that glosses are suitable as a writing and reading system neither for Deaf people, since they is not easily readable by non-linguists, nor for MT or NLG, since a computer can process information only if it is conveyed as a single input string.

## Berkeley Transcription System

The Berkeley Transcription System (BTS) was developed by Slobin *et al.* (2001) to represent utterances in ASL and NGT<sup>27</sup> in a morphological and semantic notation. It can potentially be applied to any sign language and was created to avoid SL glossing problems (Garcia and Sallandre 2013). It conveys information through a single morpheme-by-morpheme string and can provide information on various tiers, notating manual and non-manual components (including discourse markers, gaze, role shift) that can be used to create complex signs. The string is composed by words in capitals and numbers (Hoiting and Slobin 2002; Frishberg *et al.* 2012), as Figure 59 shows:

GRANDFATHER IX\_3(a) BALL CHILD IX\_3(b)-loc'INF  
^opr'NEG pm'SPHERE-src'3(a)-gol'3(b)-pth'D^

**Figure 59.** 'Grandfather gives (or gave) the child the ball.' in BTS notation  
(Frishberg *et al.* 2012: 1060)

However, Pizzuto *et al.* (2012:221) point out that BTS is not a transcription, but rather “a coding system, which shows how the coder segmented a signed production”. Furthermore, the single string modality obscures the multilinearity of sign languages.

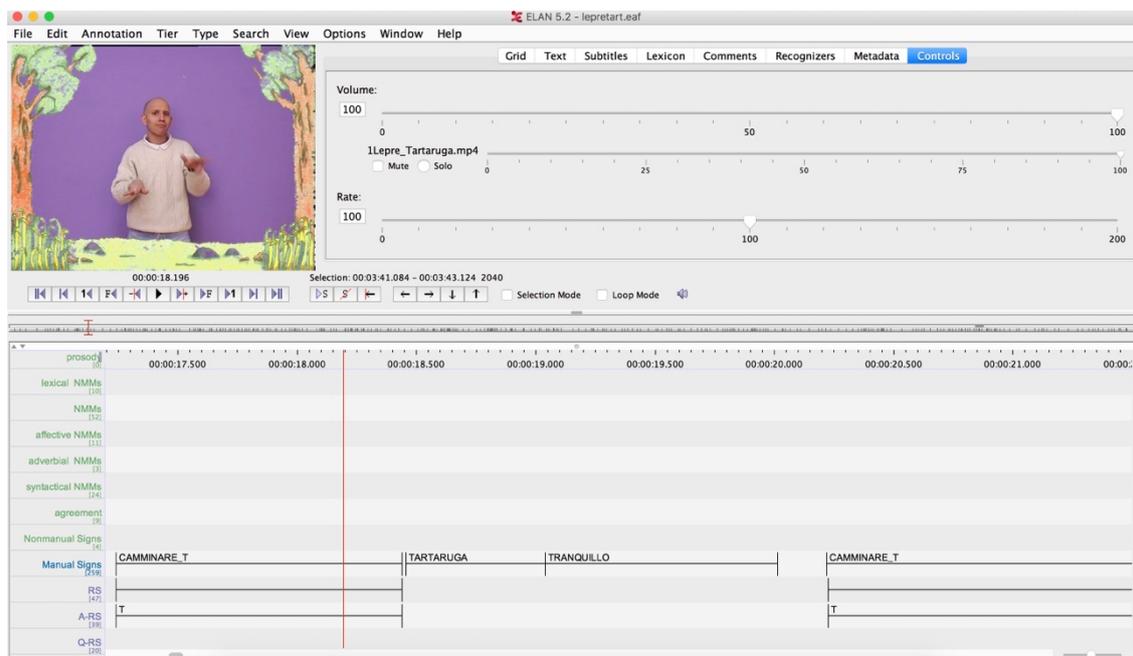
## Multimedia Language Annotation Tools

Multimedia Language Annotation Tools are tools that can be utilized to manage sign language data with a multi-level transcription. An example of annotation system is ELAN<sup>28</sup>, which I used for this work, inserting the video of the fable and manual glosses in it. It is intuitive and flexible and allows the integration of specific features into it. For instance, the Hamburg team has integrated HamNoSys into ELAN (Garcia and Sallandre 2013). The organization of the interface is showed in Figure 60 on the next page. As can be seen, videos are played on the upper left side of the window. On the upper right side users can visualize annotations in different formats such as grid, text, and subtitles. Below, on the left, there is the tier panel, which presents different tiers. The annotations can be written in the horizontal layers.

---

<sup>27</sup> (or SLN) Sign Language of the Netherlands.

<sup>28</sup> <https://tla.mpi.nl/tools/tla-tools/elan>



**Figure 60.** The ELAN interface

## 1.9 Summary

In the present chapter the reader has been introduced to the theoretical background of sign languages. I proved that sign languages are real languages and have specific characteristics. As far as Italian Sign Language (LIS) is concerned, we have seen that it has a complex phonology. I also focused on describing in details the peculiarities of this language both from a morphological and syntactic point of view. Furthermore, I explained the different functions of Non-Manual Markers, stressing the fact that they can occur in isolation. In this regard, I hypothesized that this kind of Non-Manual Markers could also appear under Role Shift, which is a peculiar strategy used in Sign Language communication. This allowed us also to acknowledge that simultaneity is a crucial peculiarity of sign languages. Finally, I introduced the main writing systems for sign languages, demonstrating that it is not easy to convert the multidimensionality and multilinearity of these languages into a single written input string.

# Chapter 2

## Processing a Natural Language

### 2.1 Introduction

In the present Chapter, Sign Languages meet Natural Language Generation (NLG) and in wider terms, Artificial Intelligence.

First, in Section 2.2, the reader will be introduced to the process of NLG and its field, which is continuously changing. I will explain how a NLG system is supposed to work and what its tasks are. The main approaches to NLG will also be illustrated. After providing this theoretical background, I will explain practical applications of NLG and analyse some existing NLG systems.

Then, we will get closer to the heart of this thesis, dedicating Section 2.3 to some systems that can generate and translate from or into sign language.

## 2.2 Natural Language Generation

### 2.2.1 What is NLG?

In order to explain what Natural Language Generation (NLG) is, I consider it necessary to take a step back and introduce first the term “Natural Language Processing” (NLP).

Studies on NLP start in the 1950s, on the basis of Chomsky’s generative grammar, the principle of the poverty of the stimulus and formal language. According to Chiari (2007), NLP is one of the most relevant fields of Computational Linguistics. It is a branch of Artificial Intelligence that deals with the interaction between humans and computers, using natural language. From a theoretical point of view, formal language plays a crucial role in this context, since NLP is based on the idea that linguistic knowledge can be described with finite formal rules.

Communication originates from the interaction between two or more participants, which are the sender and the receiver(s) or interlocutor(s): the former creates and conveys a message; the latter understands it and possibly replies. As already mentioned, this is exactly the aim of the study of NLP. On this basis, NLP can be divided into Natural Language Understanding (NLU) and Natural Language Generation (NLG).

Natural Language Understanding (NLU), or Natural Language Analysis, deals with machine reading comprehension and text analysis, namely it analyses a written or spoken input text and converts it into a formal univocal representation. However, this conversion is not always easy and this is mainly due to the interpretative ambiguity and complexity of sentences. In broad terms, we can claim that the aim of NLU is to make a machine able to *understand* a natural language, i.e. to analyse the morpho-syntactic structure of a sentence in order to realize an abstract representation of it (Chiari 2007).

Natural Language Generation exists since the 1980s (McDonald 1987) and it is the branch of NLP that deals with the generation of text in natural language, starting from an abstract representation of that language. As the reader may have noticed, the input in NLU is a well-defined object. This is not the case for NLG, since the above-mentioned representation could be a grammar, a formal system or even a knowledge base (Chiari 2007). More broadly, Gatt and Krahmer (2017) outline three general types of input for

NLG systems: non-linguistic data (data-to-text generation), texts (texts-to-texts generation) and static or moving images (vision-to-text generation).

NLG can be divided into two subtypes: elaboration and synthesis. In the elaboration phase, information used to create sentences with a specific structure is collected. The representation is then synthesised using grammatical and lexical rules of that language. In this way, the surface structure of the sentence is obtained.

Moreover, ambiguity is a crucial problem for NLU, but not for NLG. In fact, if the input representation is ambiguous, NLG systems can simply reproduce the ambiguity in the output text. Humans will then resolve the matter. However, it is essential that generators do not create ambiguity in the output if the input is not ambiguous (Gal 1991).

To sum up, NLG aims at making a machine able to *generate* grammatically correct sentences in a natural language.

As Reiter (2010) observes, the peculiarity of NLG is choice making. In fact, NLG systems have to make choices about their output texts. They can do it on the basis of linguistic correctness, such as in (1):

- (29) a) \* ‘Lana sees Lana in the mirror.’
- b) ‘Lana sees herself in the mirror.’

The author points out that in examples like (1) NLG systems can choose to use the pronoun instead of repeating the noun (binding theory)<sup>1</sup>.

A specific decision can also depend on genre constraints, for instance the repetition of a noun can be preferred in safety-critical texts, like instructions for the use of fire extinguishers, as in (2a):

- (30) a) ‘Take the fire extinguisher. Identify the pin. Pull the pin.’
- b) ‘Take the fire extinguisher. Identify the pin. Pull it.’

NLG systems may make a decision also on the basis of linguistic abilities and preferences of the reader of the text, for example they may choose (3a) instead of (3b) if the reader has not an extensive knowledge of the English language:

- (31) a) ‘I am reading a book. The book is interesting.’
- b) ‘I am reading a book. It is interesting.’

---

<sup>1</sup> For more information about the binding theory the reader is referred to Buring (2005).

### 2.2.2 The big questions for NLG

In the process of generating natural language, a NLG system must be able to decide what to say (i.e. to define a strategy), and how to say it (i.e. to define a tactic).

On one hand, it has to select the appropriate information for the task and be able to ignore irrelevant information, namely it has to define the content of the message. According to Hovy (1988), this is a very complex issue, since the decision about what to say “depends on factors about which the speaker can never have complete knowledge, such as the hearer’s knowledge and beliefs.” (Hovy 1988:153). Furthermore, the communicative goal plays a crucial role too (Reiter and Dale 2000). Ferrari (1991) suggests that this is also related to the domain and the use of the text. He claims that the semantic information that needs to be generated can be classified as follows:

- i. Definitional information, which depends on the definition of the element on which a question focuses. For example X in “What is X?”;
- ii. Descriptive information, related to the setting description. For example the description of a photo;
- iii. Historical information, which originates from time-evolving data, such as stock market indexes;
- iv. Dialogic information, related to dialogue planning and, in wider terms, to speech acts structuring.

Hovy (1988) highlights that early generators ignored this what-to-say issue and did not select and organise information, generating the whole information they received in input.

On the other hand, a NLG system has to define how to start, organise and close a message; namely it has to structure it and define its form. Thus, we deduce that syntax plays here a crucial role: first, informational units are organised with rhetorical rules and then anaphoric references and marked word orders are created. According to Ferrari (1991), rhetorical rules can differ depending on the system. For example, one could use

methods based on the discourse goal (Grimes 1975) or ATN<sup>2</sup> graphs. Furthermore, the output text could be generated with a particular format, e.g. with images, other multimedia contents or as hypertext.

“What should I say?” and “How should I say it?” are definitely two important issues that generators have to deal with. But there is another question they should also focus on: “Why should I say it?”.

As indicated by McKeown and Swartout (1988) and Hovy (1988), NLG systems should have a reason for making a decision. If they don't, they simply select the same option or a random option and, consequently, they may generate an unnatural text. For example, if NLG systems generated a text with no pronouns and only proper names in active sentences, they would produce redundant information. Furthermore, if they selected pronouns to replace proper names randomly, they would generate ambiguous information.

Decisions taken by generators can be influenced by several factors, such as the above-mentioned rhetorical rules. The role of pragmatics and the purpose of communication are not to be omitted too.

### **2.2.3 NLG tasks**

So far, we have seen that the aim of NLG systems is to generate output texts from input texts or data and that they have to make decisions on the basis of what they are supposed to say and how they say it. In order to do so, they need to perform specific tasks, which are usually six: content determination, text structuring, sentence aggregation, lexicalisation, referring expression generation, and linguistic realisation (Gatt and Krahmer 2017). I will describe them one by one in the next subsections.

---

<sup>2</sup> In wide terms, an Augmented Transition Network (ATN) is a graph structure used to store grammatical information.

## **Content determination**

As the reader may infer, content determination is the process that involves deciding what information should be communicated in a text. This task is integrated in most generators (Mellish *et al.* 2006).

According to Gatt and Krahmer (2017), the input data is usually more informative or detailed than the information we want to be conveyed in the output text. Therefore, the generator needs to filter and select data and transform them into a set of preverbal messages, which are then used in the next NLG tasks. These messages are semantic representations of information, often expressed in a formal language that distinguishes entities, concepts and relations in the domain (Reiter and Dale 1997).

Since the question of what information should be included in the output text is very application-dependent, we cannot specify general rules for this NLG task (Reiter and Dale 2000).

## **Text structuring**

Text structuring<sup>3</sup> involves deciding in which order and structure the information will be presented in the text. In fact, as Reiter and Dale (2000) observe, a text is not a random set of sentences: it has at least a beginning, a central section and an end, and it is hierarchically organised, i.e. it is composed of large constituents such as clauses made up of smaller constituents.

It is also essential to highlight that there are relations among constituents. According to Reiter and Dale (2000), they can be defined in two ways: as conceptual grouping, that means grouping information based on what they are about, or as discourse relations. Gatt and Krahmer (2017) state that discourse relations were initially expressed with handcrafted structuring rules called schemata (McKeown 1985). Later on, researchers started relying on the Rhetorical Structure Theory<sup>4</sup> (Mann and Thompson 1988; Hovy 1993).

---

<sup>3</sup> Text structuring is also known as document structuring, discourse structuring or discourse planning.

<sup>4</sup> Rhetorical Structure Theory (RST) is a theory based on tags that identify coherence relations among textual fragments.

## Sentence aggregation

Sentence aggregation involves grouping messages together in a sentence and deciding which messages have to be expressed. Although aggregation is not always necessary (Reiter and Dale 2000), it has an essential function, since it makes the output text more fluent (Cheng and Mellish 2000) and eliminates redundant information. Still, it is true that this task is not clearly defined and it is often strongly domain-dependent. Reape and Mellish (1999) prefer to distinguish between semantic aggregation at level of meaning and syntactic aggregation at level of relations among constituents. Gatt and Krahmer (2017) hypothesize that syntactic aggregation facilitates defining domain-independent rules to eliminate redundancy. They also point out that more recent approaches to aggregation tend to be data-driven: this implies that aggregation rules are directly extracted from data.

## Lexicalisation

Having determined text content and structure and grouped information together, the generator needs to choose words and phrases in order to express data in natural language. This process is called lexicalisation. It is often a complex task and this is due to one of the most fascinating peculiarities of human language: the ability to convey the same meaning in various different ways. Moreover, let us stress that lexicalisation can also involve choosing between semantically similar or related words (e.g. *bird* vs. *songbird*) and vague words denoting gradable properties (e.g. *wide* vs. *tall*) (Gatt and Krahmer 2017). Obviously, the complexity of this task depends on the generator capabilities too. In fact, once a set of alternatives is available, the system may randomly select one of them or choose a specific lexicalisation option using the choice mechanisms provided.

Lexicalisation is a highly domain- and application-dependent process as well. Pragmatics and communicative goals play also a significant role (Reiter and Dale 2000, Gatt and Krahmer 2017). For instance, the list in (4) is awkward and redundant:

- (32) a. “There was rain on every day for eight days from the 11th to the 18th.”
- b. “There was rain on every day for eight days from the 11th.”
- c. “There was rain on every day from the 11th to the 18th.”

d. “There was rain on the 11th, 12th, 13th, 14th, 15th, 16th, 17th, and 18th.”

(Reiter and Dale 2000:54)

We usually do not utter such a discourse. However, it may be appropriate when expressed in a weather report, if the goal is to stress that it has rained non-stop for a certain period of time.

In light of the above, the fact that the process of lexicalisation has also been studied from a psycholinguistic point of view (Levelt 1989; Levelt *et al.* 1999) is not surprising.

### **Referring expression generation**

Referring Expression Generation (REG) consists in the actual selection of words or phrases in order to identify domain entities (Reiter and Dale 1997). As the reader may have inferred, this task is closely related to lexicalisation. However, REG is a discrimination process, where the system produces a description of an entity that allows identifying that entity in a specific context (Reiter and Dale 2000). The authors observe that two types of references exist. The former is the initial reference, which allows the identification of an entity for the first time and establishes the discourse perspective. The latter is the subsequent reference, which identifies an entity that has already been mentioned in the discourse and distinguishes it from other referents. Hence, the initial reference is related to referential form, that is the form used to refer to the referent, e.g. a pronoun, a proper name or a description (Gatt and Krahmer 2017). Moreover, if the chosen form is a description, it is necessary to determine the referential content, i.e. to describe the entity’s properties in order to identify it.

Much work has been devoted to this task in recent years (Mellish *et al.* 2006; Siddharthan 2011) and also to the interface between REG and computer vision to create descriptions of objects in visual scenes (Mitchell *et al.* 2013; Kazemzadeh *et al.* 2014; Mao *et al.* 2016).

### **Linguistic realisation**

Last but not least, the generator needs to combine words and phrases to produce a well-formed text. This task is referred to as linguistic realisation and involves the use of grammatical rules that govern morphology and syntax. For example, from a morphological point of view, the system has to produce the appropriate verb

conjugations and agreement, if needed. From a syntactic point of view, it has also to order constituents correctly and insert function words such as auxiliary verbs and prepositions. Moreover, let us not forget that the generator has to be able to produce an orthographically correct text too (Reiter and Dale 1997, 2000).

Gatt and Krahmer (2017) suggest that many approaches can be adopted to produce an output, for instance the use of human-crafted templates, hand-coded grammar-based systems or statistical approaches.

Templates are useful “when application domains are small and variation is expected to be minimal” (Gatt and Krahmer 2017:19). An example of template is shown in (5a) and can generate the sentence in (5b):

(33) a) \$player scored for \$team in the \$minute minute.

b) Ivan Rakitic scored for Barcelona in the 4th minute.

(Gatt and Krahmer 2017:19)

This approach allows avoiding ungrammatical structures, but it is very laborious if templates are constructed by hand.

Hand-coded grammar-based systems make decision on the basis of the grammar of the language on which they are working. An advantage of these systems is that they are domain-independent. Nevertheless, they are often not able to choose among related options (Gatt and Krahmer), as in (6):

(34) a) Ivan Rakitic scored for Barcelona in the 4th minute.

b) For Barcelona, Ivan Rakitic scored in minute four.

c) Barcelona player Ivan Rakitic scored after four minutes.

(Gatt and Krahmer 2017:19-20)

Other approaches described by the authors are statistical approaches, which are increasingly used. Some of them involve the use of statistical information to filter the output and of a handcrafted base generator (see Langkilde-Geary 2000; Langkilde-Geary and Knight 2002). Other methods utilise exclusively statistical information (see Espinosa *et al.* 2008; White and Rajkumar 2012).

This subdivision denotes a traditional, modular architecture of an NLG system. However, as Reiter e Dale (1997) point out, the six-task distribution does not imply a six-module architecture of the generator. In fact, the architecture of many systems is composed by one module only, where that module performs tasks<sup>5</sup>.

Let us now see what the main approaches to NLG are.

#### **2.2.4 NLG approaches**

In the current state of the art of NLG we can find three main kind of approaches: rule-based, planning-based and data driven approaches (Gatt and Krahmer 2017).

Rule-based or modular approaches are traditional approaches used in the early Artificial Intelligence research. They provide for a three-module architecture of the NLG system that can be summarised as follows: first decide what to say, then decide how to say it and finally say it.

The first module is the Text Planner<sup>6</sup>, which decides what to say, carrying out two tasks. In fact, it determines what information is relevant (content determination) and how it should be structured (text structuring), based on discourse goals and a knowledge base. The output of this module is usually a text plan, which is a tree structure of messages that consist of information extracted from the input data.

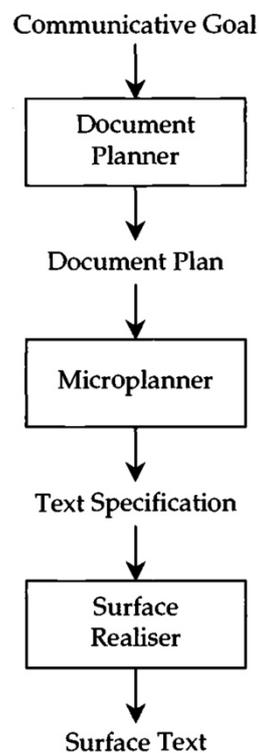
The text plan becomes then the input for the second module: the Sentence Planner (or Microplanner), which decides how the information should be linguistically expressed. More precisely, it groups messages together in a sentence (aggregation), it chooses content words in order to express domain data (lexicalisation) and it chooses referring expressions to identify referents (referring expression generation). The output of the Sentence Planner may be a text or a text specification, i.e. a tree that specifies the text inner structure (e.g. paragraphs) and syntactic structures of sentences (Reiter and Dale 2000; Reiter 2010).

---

<sup>5</sup> The reader is referred to Appelt (1985) for an example of such systems.

<sup>6</sup> Also known as Document Planner, Macroplanner or Text Planning Component.

At this point, the generator solely needs to generate an actual text, which has to be morphologically, syntactically and orthographically correct. This task is performed by the third module, which is the Surface Realiser<sup>7</sup> (Reiter and Dale 2000; Reiter 2010; Gatt and Krahmer 2017). The traditional architecture of NLG systems is shown in Figure 1.



**Figure 1.** Traditional NLG system architecture diagram  
(Reiter and Dale 2000:60)

As the reader may have noticed, this approach shows well-defined boundaries between tasks, modules and especially between what to say and how to say it. However, many authors, such as Mellish *et al.* (2006), point out that in reality these divisions are blurred and the pipeline may vary from system to system (Gatt and Krahmer 2017). For example, text structuring and aggregation, that deal with the structure of text and sentence respectively, may not necessarily have to be carried out by two different modules.

This approach has also been challenged by planning-based and data-driven approaches that tend to blur the divisions between modules.

---

<sup>7</sup> Also known as Linguistic Realiser or Surface Realisation Component.

Planning-based perspectives are based on the planning paradigm: an action such as a task's completion that is performed to achieve a communicative goal produces a new state, i.e. a change in context. As Gatt and Krahmer (2017:26) observe, "there is in principle no restriction on what types of actions can be incorporated in a plan"; therefore such an approach does not impose well-defined boundaries, but rather a more integrated architecture. Nonetheless, it does not allow expressive formalisms to represent constraints in NLG. Hence, a system cannot reason about interlocutor's beliefs and intentions.

In recent years, a different kind of approach has been frequently employed: data-driven approaches. Data-driven are global, integrated approaches and are grounded on statistical learning of correspondences between input data and output text and allow unified formalisms. However, when adopting this perspective, a specific problem arises, namely the acquisition of the right data. Indeed, researchers cannot actually predict whether existing data-text alignment methods are suitable for large amounts of heterogeneous data or not (Gatt and Krahmer 2017).

As the reader may have inferred, NLG is a very complex field that can potentially be applied in several context. I will discuss this topic in the next section.

### **2.2.5 NLG applications**

As we now know, NLG systems must operate making decisions on different levels and generation is influenced by the purpose for which text is generated. Therefore, it is not surprising that NLG finds use in different sectors. Initially, NLG underlay question-answering systems. Later on, interacting systems have been built, such as chess-playing systems, which use planning in order to make the most appropriate move. In time, systems have become increasingly complex and sophisticated. They are now able to operate in various fields and generate recipes, fables, tales and different descriptions, e.g. geographical descriptions.

Some examples of NLG applications are shown on the next page.

- i. Reports of various kinds:
  - Weather reports (Goldberg *et al.* 1994; Coch 1998; Reiter *et al.* 2005; Turner *et al.* 2008; Ramos-Soto *et al.* 2015),
  - Financial reports (Kukich 1983; Plachouras *et al.* 2016),
  - Sport reports (Robin and McKeown 1996; Chen and Mooney 2008),
  - Medical documents and reports (Hüske-Kraus 2003, Gatt *et al.* 2009; Portet *et al.* 2009),
  - Engineering reports (Yu *et al.* 2007);
- ii. Virtual newspapers from sensor data (Molina *et al.* 2011);
- iii. Text intended to persuade (e.g. encouraging people to stop smoking – Reiter *et al.* 1999) or motivate users (Reiter *et al.* 2003), make them less anxious (Cawsey *et al.* 2000) or entertain them (Binstead and Ritchie 1997);
- iv. Drafts of various documents, such as:
  - Instruction manuals (Paris *et al.* 1995),
  - Legal documents (Sheremetyeva *et al.* 1996);
- ii. Simplification of complex texts (Siddharthan 2014; Macdonald and Siddharthan 2016);
- iii. Automatic spelling and grammar correction (Dale *et al.* 2012);
- iv. Automatic generation of questions for educational and other purposes (Brown *et al.* 2005; Rus *et al.* 2010);
- v. Descriptions of static or moving images based on computer vision input (Kulkarni *et al.* 2013; Thomason *et al.* 2014);
- vi. Supporting users with disabilities, e.g. helping blind people access graphical information (Ferres *et al.* 2006), and helping non-speaking users tell stories (Reiter *et al.* 2009);
- vii. Supporting deaf people to learn the syntax of written English (McCoy *et al.* 1996);
- viii. Generating jokes (Binstead and Ritchie 1997).

## 2.2.6 Some examples of NLG systems

Let us now try to understand how generators work. In this Section I will describe five NLG systems, which operate in the fields of weather forecasting, smoking-cessation persuasion, marketing, entertainment and education respectively: FoG, STOP, DYD, JAPE and ICICLE.

### FoG

FoG (Forecast Generator – Goldberg *et al.* 1994) is one of the first landmark NLG systems in the domain of weather reports. Developed by the NLG software house CoGenTex for the Canadian weather service, it generates weather reports in English and French from numerical weather simulations, deciding how detailed they have to be. The input information consists of predictions about meteorological phenomena like wind speeds and rainfall in a specific region in a specified time interval and it is then elaborated and summarized in an output text, as shown in Figures 2 and 3 (Reiter and Dale 2000).

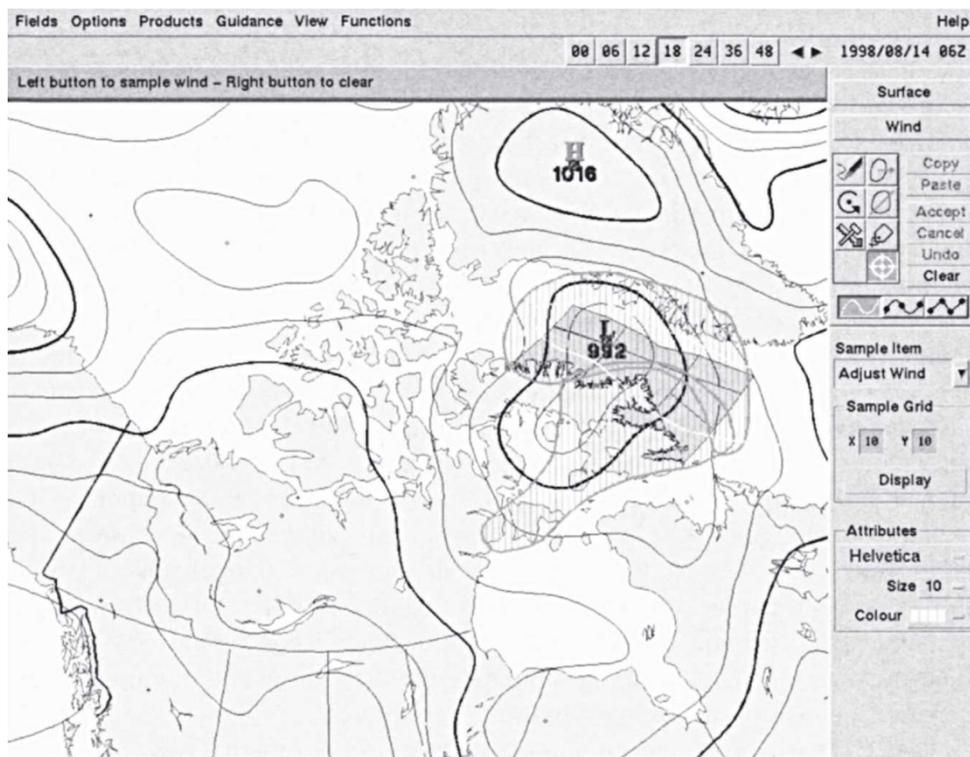


Figure 2. An example of FoG input (Reiter and Dale 2000:11)

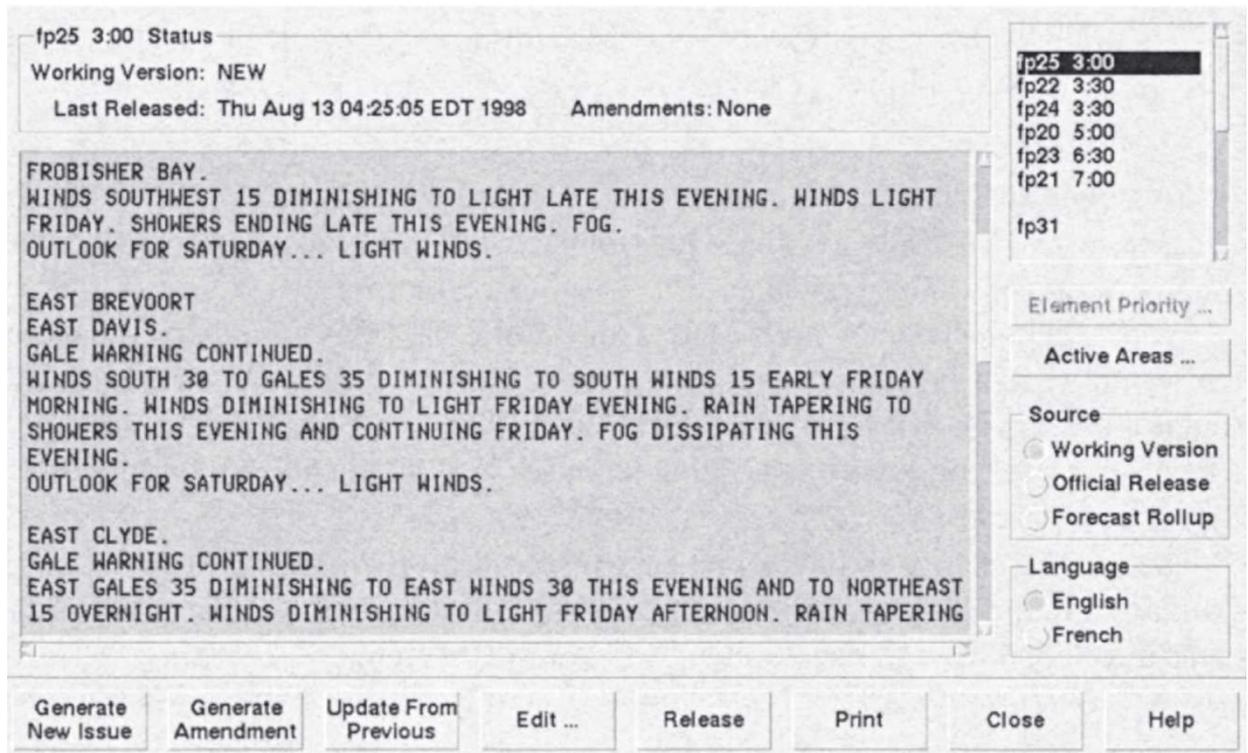


Figure 3. An example of FoG output. (Reiter and Dale 2000:11)

## STOP

The STOP (Smoking Termination through cOmputerised Personalisation) system generates short personalised smoking-cessation letters. Personalisation is based on responses to a four-pages questionnaire that smokers fill out and that asks about their smoking habits and beliefs, health problems, attempts to quit and so on (Reiter *et al.* 1999). For this reason, the STOP system is defined as a “paper-in, paper-out” system: “users fill out a paper questionnaire, and receive in response a paper leaflet” (Reiter *et al.* 1999:3). It is based on various types of knowledge, e.g. psychological knowledge about addictive behaviours and medical knowledge about smoking and categories of smokers, from people who want to quit smoking to the ones that are not willing to. An example of page generated by STOP is shown in Figure 4 on the next page.

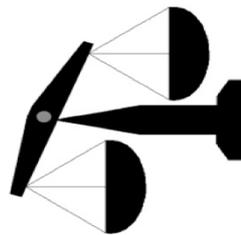
# Smoking Information for Fiona Cameron

## You have good reasons to stop...

People stop smoking when they really want to stop. It is encouraging that you have many good reasons for stopping. The scales show the good and bad things about smoking for you. They are tipped in your favour.

### THINGS YOU LIKE

you enjoy it  
it's relaxing  
it stops stress  
it relieves boredom



### THINGS YOU DISLIKE

it's bad for you  
it makes you less fit  
it's a bad example for kids  
you're addicted  
it's unpleasant for others  
it's a smelly habit  
it's expensive  
it's bad for others' health

## You could do it...

Most people who really want to stop eventually succeed. In fact, 10 million people in Britain have stopped smoking - and stayed stopped - in the last 15 years. Many of them found it much easier than they expected.

You are right to think that if you tried to stop smoking you would have a good chance of succeeding. You have several things in your favour.

- You have stopped before for over a year.
- You are a light smoker.
- You have good reasons for stopping smoking.
- You expect support from your partner, your family, your friends, and your workmates.

We know that all of these make it more likely that you will be able to stop. Most people who stop smoking for good have more than one attempt.

## Overcoming your barriers to stopping...

You said in your questionnaire that you might find it difficult to stop because smoking helps you cope with stress. Many people think that cigarettes help them cope with stress. However, taking a cigarette only makes you feel better for a short while. Most ex-smokers feel calmer and more in control than they did when they were smoking. There are some ideas about coping with stress on the back page of this leaflet.

You also said that you might find it difficult to stop because you would *get bored*. It's a habit to smoke when you have nothing to do. If you decide to stop it might be worth planning how you could keep yourself busy.

## And finally...

We hope this letter will help you feel more confident about giving up cigarettes. You know why you'd like to stop, you just need to decide to do it! If you have a go, you have a real chance of succeeding.

With best wishes,

Great Western Road Medical Group.



Figure 4. An example of page generated from STOP.

(Reiter *et al.* 1999:4)

## DYD

The DYD (Dial-Your-Disc) system (van Deemter and Odijk 1997) is an interactive system that produces spoken information derived from information stored in a database about a CD of Mozart's musical compositions. More precisely, the spoken monologue is generated if the user selects a track. Users can communicate what they want to hear in written or spoken English, in abbreviated form. For example, if they communicate the term "sonata", DYD selects the set of tracks that are part of a sonata, as shown in Figure 5. This kind of spoken description is intended to increase interest in that CD. Furthermore, it could be used for teleshopping or even for educational purposes (van Deemter and Odijk 1997).

**User request (1):** *sonata.*

**System response:** *You asked for a sample in the sonata genre. I found 163 tracks. You will hear an example.*

*The following+ composition+, from which you are going to hear a fragment+ of part one+, was written+ by Mozart in his Salzburg+ period, in Salzburg, in Brussels+, and in Paris+. The KV+ number of this work+ in c+ is K. six+. It is also+ known as the first+ sonata+ for violin+, and harpsichord+. The violin+ is played by Gerard Poulet+, and the harpsichord+ by Blandine Verlet+. The violin+ sonata consists of four+ parts+: allegro+, andante+, minuet tempo di menuetto+ - tempo di menuetto, and allegro molto+. This allegro is located on the first+ track+ of the second+ CD+ of volume fifteen+. This allegro lasts four+ minutes+ twenty+ nine+ seconds+. The recording+ of this sonata+ was made in Paris+, France+, in the seventies+. The quality+ of the recording of this work is ADD+. The following+ is a fragment+ of this allegro+. (There follows a fragment.)*

**Shortcomings:** (a) The *from which* construction is slightly awkward, but can easily be replaced. (b) The second occurrence of *this allegro* is infelicitous. (*It* would have been better.) (c) It seems clearly inappropriate that the second occurrence of *Salzburg* is deaccented. (d) Likewise, the second occurrence of *tempo di menuetto* could better be accented.

**Figure 5.** An example of DYD output. (van Deemter and Odijk 1997:27)

## **JAPE**

JAPE (Joke Analysis and Production Engine) (Binstead and Ritchie 1997, Ritchie 2003) is a punning riddles generator. It produces text strings, using a set of syntactic categories and a lexicon. Some examples of produced riddles are shown in (7) and (8):

(35) “What is the difference between leaves and a car? One you brush and rake, the other you rush and brake.”

(36) “What do you call a strange market? A bizarre bazaar.”

(Ritchie 2003:2)

It is based on four different types of rules:

- i. Schemata, which define configuration of lexemes underlying puns;
- ii. Sentence forms, which are canned sentences with slots for further text to be inserted;
- iii. Templates, which define conditions for specific terms to be inserted in sentence forms;
- iv. Small-Adequate-Description generation rules, which create abstract representations from lexemes (Ritchie 2003).

## **ICICLE**

The ICICLE (Interactive Computer Identification and Correction of Language Errors) system (McCoy *et al.* 1996) was created to help deaf people learn the syntax of written English. The input information consists of the text typed by the user, whose sentences are analysed one by one by the Error Identification component. Once errors are identified, the Response Generator has to decide how to correct them. The output response is then presented to the user, who can correct his or her own text and have it re-checked (McCoy *et al.* 1996).

An example of ICICLE User Interface is shown in Figure 6 on the next page.

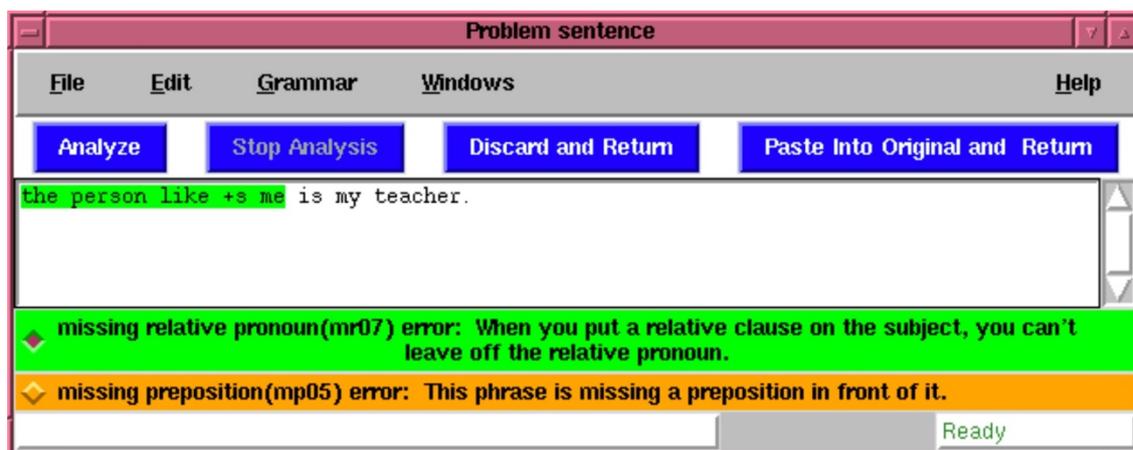


Figure 6. An Example of ICICLE User Interface. (Michaud *et al.* 2000)

If the reader is interested in seeing first-hand how a NLG system works, some generators are available for download in the Web page of the Association for Computational Linguistics.<sup>8</sup>

## 2.3 From words to signs, from signs to words

### 2.3.1 Foreword

In Section 2.2, I provided an overview on Natural Language Generation at the service of humans. In recent years, research in the field of NLG, and more generally of NLP and Artificial Intelligence (AI), has progressed and has further investigated natural vocal languages. However, as we now know, natural languages are not exclusively vocal: sign languages are fully formed natural languages too.

The first pioneering studies on sign languages have generated a growing interest in Deaf communities and its languages, from different points of view. In this respect, in the last thirty years several technology tools have been developed with the aim of helping Deaf people and making the hearing world more accessible to the Deaf world, and vice versa. In other words, AI has extended its scope to sign languages too. Hence, it is now

<sup>8</sup> [https://aclweb.org/aclwiki/Downloadable\\_NLG\\_systems](https://aclweb.org/aclwiki/Downloadable_NLG_systems)

possible to find tools like the app StorySign<sup>9</sup>, which helps Deaf children to read written stories in vocal language through an avatar that uses sign language<sup>10</sup>. In this context, an avatar is a customisable virtual character (usually with human form) that can be displayed on a monitor or a screen. The avatar used by StorySign is shown in Figure 7.



**Figure 7.** Star, the avatar used by StorySign. (screenshot of a video<sup>11</sup>)

Furthermore, researchers have been working on creating machines that generate both from sign and vocal language. Actually, much work has been done in the field of Machine Translation (MT), which is a sub-field of computational linguistics that deals with translation of text or speech from an actual natural language to another.<sup>12</sup>

I could write reams about this topic (and to be honest I really wouldn't mind doing it), but this is not the aim of my thesis. For now, all the reader needs to know is that, as for NLG, there are different approaches to MT, for example statistical and rule-based approaches.

Statistical MT relies on statistical models and bilingual text corpora, which can be aligned in order to infer syntactic rules or simply to extract phrases.

Rule-based approaches can be divided into direct, transfer and interlingua approaches. A direct translation is the literal translation of words and phrases. Transfer-based translations analyse the source text and create a syntactic representation of it, which is then converted to an appropriate representation in the target language from which the

---

<sup>9</sup> <https://consumer.huawei.com/en/campaign/storysign>

<sup>10</sup> For further information, please see <https://www.aardman.com/blog/creating-storysign-app-for-deaf-readers> and <https://www.youtube.com/watch?v=rQazllUknIU>.

<sup>11</sup> <https://youtu.be/rQazllUknIU?t=134>

<sup>12</sup> The reader is referred to Gal (1991) for further reading.

translated text is generated. The interlingua approach allows the deepest analysis of the source text, which is converted into a language-independent representation that preserves text meaning and that is the input for the generation. These levels of analysis are illustrated in the Vauquois Triangle in Figure 8:

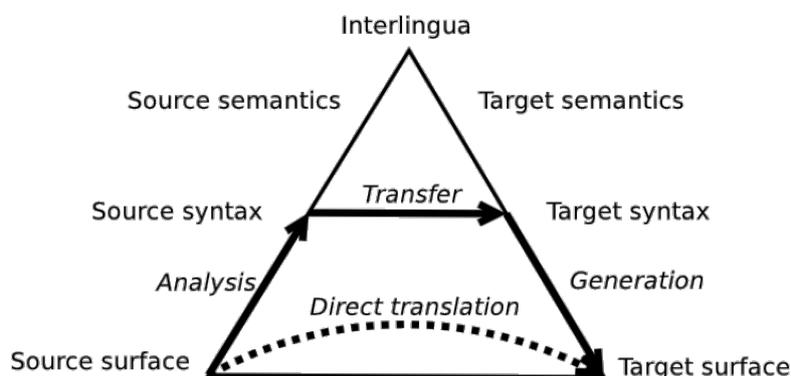


Figure 8. The Vauquois Triangle. <sup>(13)</sup>

As far as sign languages are concerned, researchers have tried to build MT systems that are able to translate from vocal and sign languages. Early research mostly adopted rule-based approaches (Stein 2012). This is also due to the fact that statistical approaches require a large amount of data, which were initially not available.

However, I would like to stress that the structure of a system that translate from vocal to sign language is different from the structure of a system that translate from sign to vocal language. It is quite easy to understand why. In fact, the former needs an avatar, or rather an animation system, to perform the output in sign language; the latter does not, but needs a recognition system to recognise the input in sign language.

The next sections (§ 2.3.2 and § 2.3.3) are devoted to providing an overview of MT systems in Sign Language (SL) field. For the sake of brevity, I will only provide an overview of some systems, so that the reader may get a rough idea of how such machines work.

<sup>13</sup> <http://mttalks.ufal.ms.mff.cuni.cz/index.php?title=Intro>

### 2.3.2 From vocal to sign language

Much MT work in SL field focuses on generation and translation from vocal language to sign language. In this Section, I will illustrate how such systems work. More precisely, I will focus my attention on Zardoz, ViSiCAST, the system built by Wu *et al.* (2007) and the Italian ATLAS and LIS4ALL.

#### Zardoz

One of the first attempts in this sense was Zardoz (Veale and Conway 1994; Veale *et al.* 1998), an interlingua-based translation system that translates English speech and text into animated Sign Language (more specifically in American, Irish and Japanese Sign Language) using a blackboard control structure (Cunningham & Veale, 1991; Veale & Cunningham, 1992). It is a cross-modal system and its architecture can be divided into different panels, as shown in Figure 9:

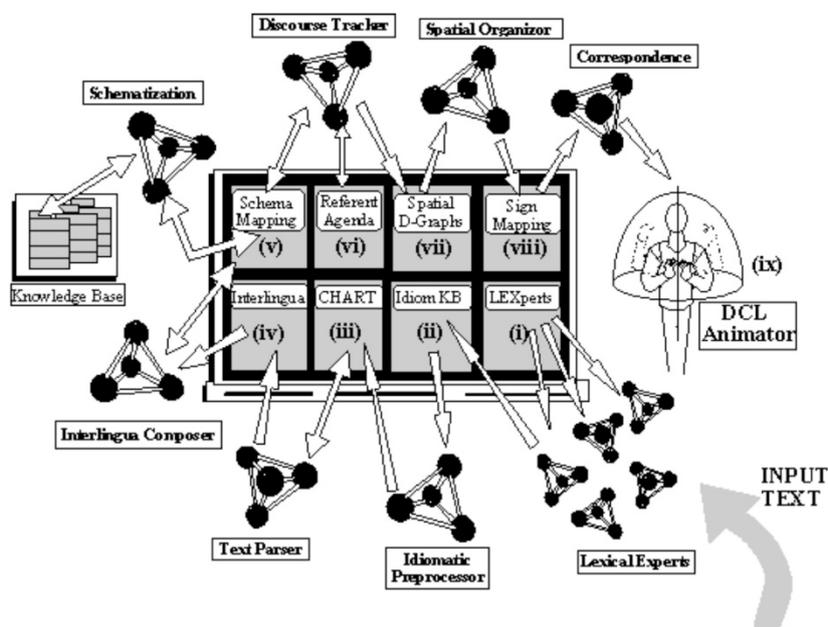


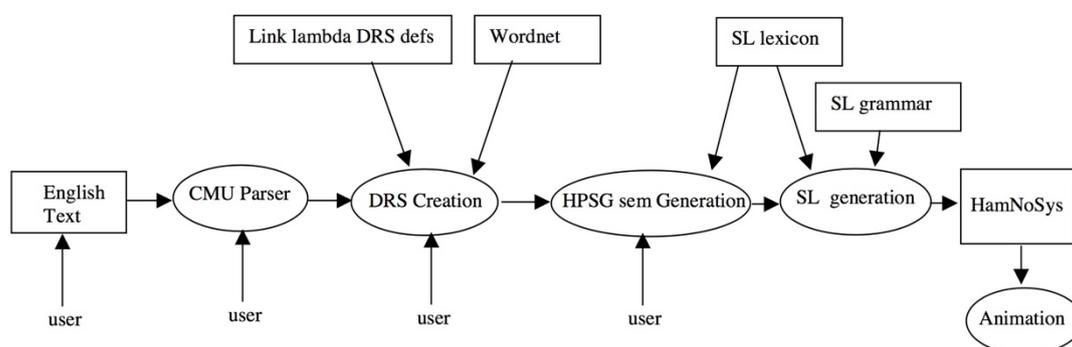
Figure 9. The Architecture of Zardoz. (Veale *et al.* 1998:86)

As we can see, the input text is lexically analysed (i), idiomatically reduced (ii) and parsed (iii) using a unification grammar, in order to create a deep syntactic and semantic representation. The representation is transformed into a glossed interlingua representation (iv). Then, the system removes metaphors and metonyms specific to the source language, through a schematisation process (v), in order to produce a language-

independent representation. Subsequently, the discourse tracker creates the appropriate anaphoric relations (vi) and the sign-syntax agency generates the SL linear order, using a scheme of spatial dependency graphs (vii). Finally, the sign-mapping agency creates correspondences between concepts and signs (viii). In this way, the interlingua structure transforms into a stream of sign tokens, which are compiled into the Doll Control Language program. This program animates an avatar and produces the output (ix). The avatar can realise both manual signs and non-manual components.

## ViSiCAST

The ViSiCAST (Virtual Signing, Animation, Capture, Storage and Transmission) Project (Sáfár and Marshall 2001; Marshall and Sáfár 2003) is a three-year project (2000-2002) funded by the European Community within the Seventh Framework Programme. The project develops an interlingua system with the aim of translating from English text to Sign Language of the Netherlands (NGT), German Sign Language (DGS) and British Sign Language (BSL). The system produces a representation on the basis of Discourse Representation Structures (DRSs) and converts them into sign languages (Elliott *et al.* 2000). For example, as far as translation from English to BSL is concerned, it displays four main translation stages: English syntactic parsing, DRSs creation, semantic transfer and generation of HamNoSys phonetic descriptions. The overall architecture is shown in Figure 10:



**Figure 10.** Architecture of the MT system used for the ViSiCAST project.

(Marshall and Sáfár 2003:114)

As the reader can see in Figure 10, some components are augmented by-user interaction. The input for the system consists in an English text, which is then parsed by the CMU (Carnegie Mellon University) link grammar parser to create syntactic

dependencies. If more than one dependency is possible, the user can manually select the appropriate one. Then, an intermediate representation, i.e. the DRS, is built in order to capture semantic content and resolve anaphora. Thus, in this stage an English-oriented DRS is converted into a sign language-oriented DRS, which is then transformed into a HPSG (Head-Driven Phrase Structure Grammar) semantic structure. At this point, SL generation begins, on the basis of sign language lexicon and grammar. The output of this stage is a HamNoSys description of a sign sequence, which is finally realised by the human avatar. As Marshall and Sáfár (2003) point out, the realisation of Non-Manual Components is not provided. However, they claim that it will be taken into account for future research.

### **The system of Wu *et al.* (2007)**

Wu *et al.* (2007) present a transfer-based statistical MT model, which translates Chinese into Taiwanese Sign Language (TSL) in three stages, using Probabilistic Context Free Grammars (PCFG).

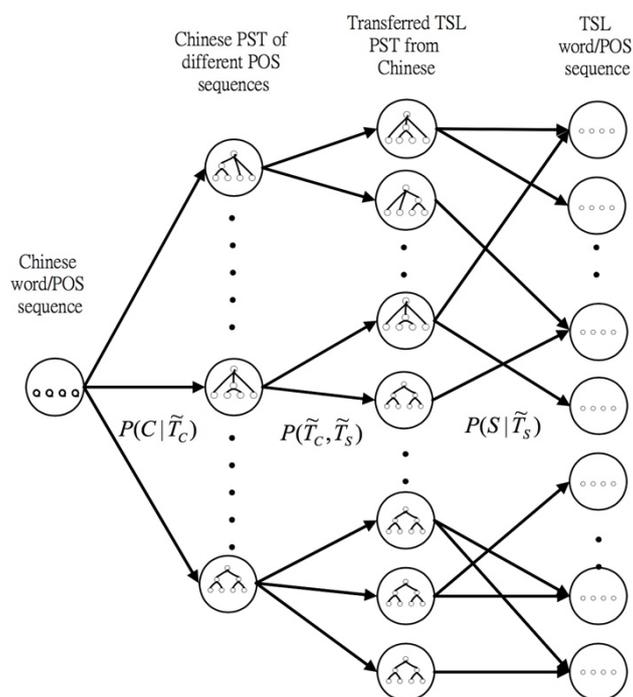
First of all, the input Chinese text is segmented into words, which are parsed into possible Phrase Structure Trees (PSTs), using the PCFG of the Chinese Treebank.

Then, transfer probabilities deriving from the Chinese-TSL corpus are employed to transfer CFG rules to the corresponding sign language PSTs.

Finally, these PSTs produce the TSL output and the Viterbi algorithm is used to refine the translation (Wu *et al.* 2007).

The above-mentioned stages are illustrated in Figure 11 on the next page.

As far as Italian Sign Language is concerned, there are some systems that work in this direction too: ATLAS and LIS4ALL.



**Figure 11.** The three-stage translation (Wu *et al.* 2007:8)

## ATLAS

The aim of the ATLAS (Automatic Translation into Sign Language) project (2010-2012) was to build a system able to translate written Italian into LIS through a virtual avatar, in order to make accessible media content to Deaf people, starting from the weather forecasting domain. It relies on a statistical and a rule-based translator.

The system transforms the input text into an ontology-based logical representation, which is then processed by a linguistic generator that generates glossed LIS sentences. These are defined through the AWLIS (Atlas Written LIS) formalism and are sent to the animation module. This module consists of a Planner that sets signs into the signing space and an Animation Interpreter, which produces LIS signs and facial expressions stored in the dictionary called “signary” (Barberis *et al.* 2011; Lombardo *et al.* 2011).

The architecture of this system is shown in Figure 12 on the next page.

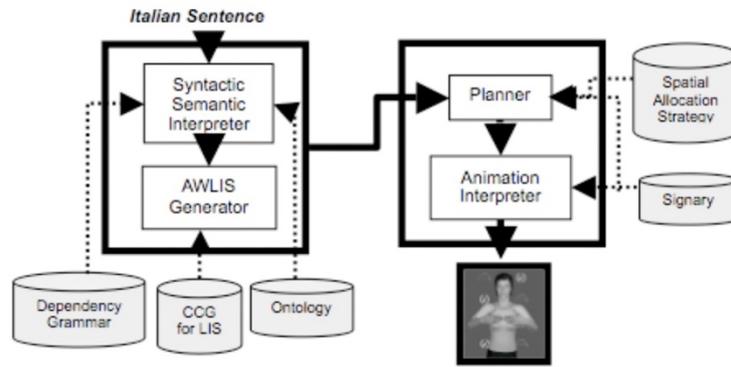


Figure 12. Architecture of the ATLAS system. (Barberis *et al.* 2011:3)

### LIS4ALL

LIS4ALL focuses on translating Italian railway train announcements in LIS. The system includes four modules: a regular expression-based analyser that produces the semantics to parse the Italian input, a semantic interpreter, a generator and an avatar performing the final LIS signed sentence. The generator is composed of a Microplanner and a Realiser. As we know, the Microplanner decides how to express information. Hence, it decides which signs to use. Then, it generates a tree structure, which becomes the input for the Realiser. This second sub-module produces sentences in AWLIS, as in the ATLAS project, and sends them to the avatar (Battaglino *et al.* 2015), as shown in Figure 13.

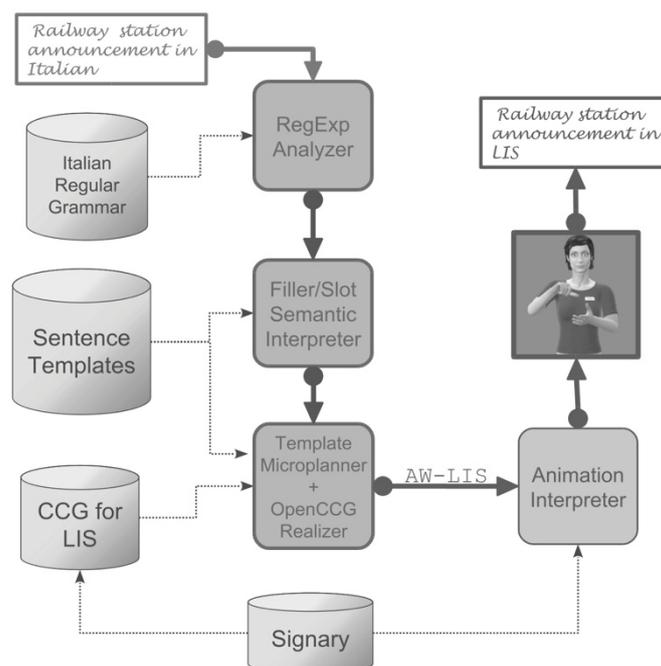


Figure 13. LIS4ALL Architecture. (Battaglino *et al.* 2015:341)

The reader who is interested in other vocal to sign language systems can consult the following references: Grieve-Smith (1999) (Albuquerque Weather – English to BSL); Zhao *et al.* (2000) (TEAM Project – English to ASL); Huenerfauth (2006) (Multipath Approach – English to ASL); Stroppa and Way (2006), Morrissey (2008), Morrissey (2011), Morrissey and Way (2013) (MaTrEx – English to ISL<sup>14</sup> and vice versa).

For further reading on this topic, the reader is also referred to Stein *et al.* (2012), Ebling (2016) and Schmidt (2016), to mention a few.

Personally, I think that systems that translate from a vocal to a sign language could be extremely useful, especially in some circumstances, when a human interpreter cannot be physically present or his/her “participation” to the context would not be somehow sensible.

Let us take for example the case of LIS4ALL. Let us try to imagine the following: an avatar displayed on a monitor in a railway station informs using sign language that the train that was expected to arrive at platform 1 is now arriving at platform 14. Or even, an avatar displayed on a monitor in a train informs in sign language that that train is going to reach its destination 40 minutes late due to a technical problem.

Accessibility. It would be a big step forward.

### **2.3.3 From sign to vocal language**

As we have seen, there are many systems that strive to translate from a vocal language to a sign language. However, researchers have tried to move to the opposite direction too, with translators from sign to vocal language. In order to create a complete MT system of this kind, much research has focussed on visual recognition, especially on recognition of Non-Manual Markers. Indeed, since sign languages are conveyed through the visual channel, experts have tried to develop systems that are able first of all to recognise a visual input and finally to generate an output in vocal language. In the next subsections, I will introduce the following sign to vocal language translation and generation systems: the system of Bauer *et al.* (1999), SignSpeak and LSESpeak.

---

<sup>14</sup> Irish Sign Language.

## The system of Bauer *et al.* (1999)

Bauer *et al.* (1999) aim at translating German Sign Language (DGS) into German text within the domain of “shopping at the supermarket”, building a SL MT system composed of two main modules: a video-based continuous sign language recognition tool and a statistical translation tool, as illustrated in Figure 14.

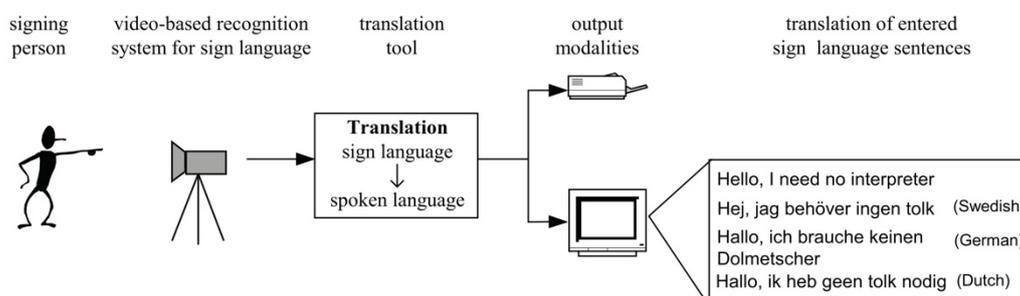


Figure 14. Components of the system. (Bauer *et al.* 1999:4)

First of all, the input video is recorded using a video camera and it is segmented. Next, features of sign parameters are computed and converted into feature vectors, which are the input for the SL recognition system. Every sign is then classified using Hidden Markov Models<sup>15</sup> (Rabiner and Juang 1989). After the classification of the sentence, signs are processed by the translation tool, which converts them into words of a meaningful sentence in German. Moreover, Bauer *et al.* (1999) indicate that this tool consists of a language model and a translation model that can be divided into a lexical and an alignment model.

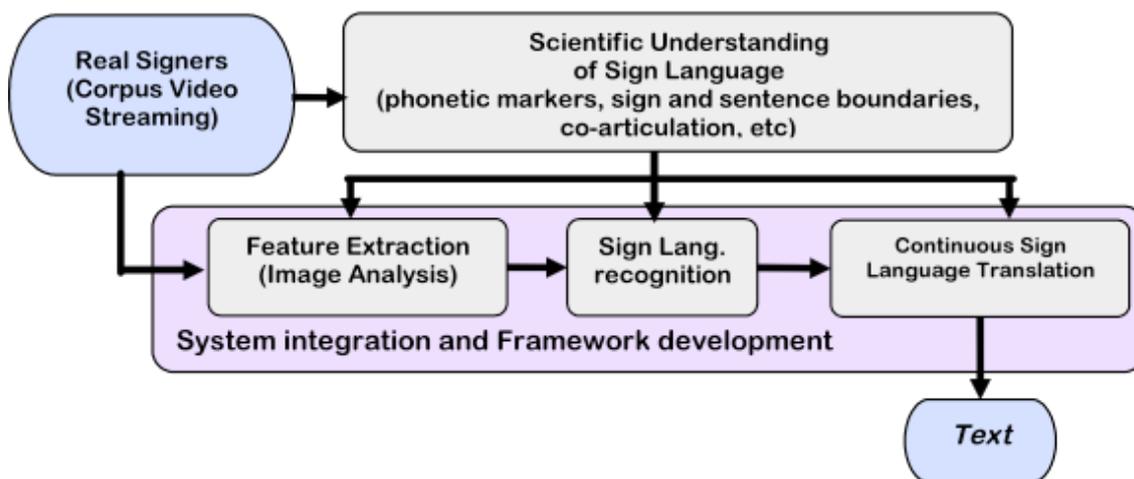
## SignSpeak

SignSpeak is a three-year project (2009-2012) financed by the European Community within the Seventh Framework Programme. Its aim is to develop a statistical machine translation system that is able to recognize and translate continuous sign language into text (Dreuw *et al.* 2010). More precisely, it translates NGT and DGS in Dutch and German respectively and it is based on videos of the Corpus NGT and of some DGS databases (<sup>16</sup>).

<sup>15</sup> A Hidden Markov Model is a statistical model in which states are unobservable.

<sup>16</sup> <https://www.ru.nl/sign-lang/projects/completed-projects/signspeak/>

The system is composed of two main modules: a sign language recognition system and a sign language translation system. The former converts the visual information into an intermediate representation of sign language utterances, which is expressed with glosses. The latter translates this representation into a text in vocal language (Schmidt 2016). Figure 15 illustrates a conceptual scheme of the work planned within this project:

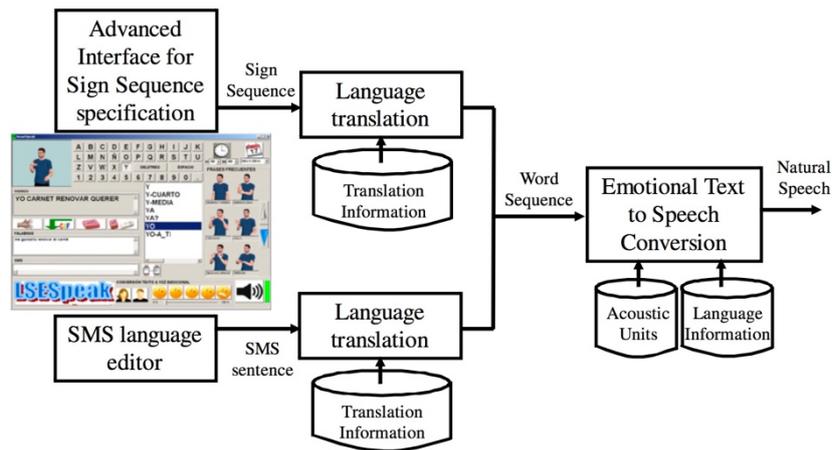


**Figure 15.** Conceptual scheme of the work planned in the SignSpeak project. <sup>(17)</sup>

## **LSESpeak**

LSESpeak (López-Ludeña *et al.* 2013) is a system that generates spoken Spanish from Spanish Sign Language (LSE). It does not provide for a video-based recognition system, but consists of two different tools: a “LSE into Spanish translator system” and an “SMS to Spanish translator system”. They are both composed of three modules: an interface that allows the user to indicate a sequence of signs in LSE or an SMS message, a translator that converts LSE or the SMS message into written Spanish, and an emotional text-to-speech converter. The converter enables the user to produce a speech message with specific characteristics, choosing the voice gender (male or female), the emotional state (e.g. happy, sad) and the emotional strength. The module diagram of LSESpeak is shown in Figure 16 on the next page.

<sup>17</sup> <http://www.signspeak.eu/en/overall.html>



**Figure 16.** Architecture of LSESpeak (López-Ludeña *et al.* 2013:1285)

Let us focus on the translation from LSE task. The LSE to Spanish translation module is phrase-based and has been trained on parallel corpora. First of all, a statistical machine translation toolkit called GIZA++ aligns Spanish words and LSE signs. Then, phrases that are consistent with the alignment are extracted. In the last stage, i.e. phrase scoring, translation probabilities are computed (López-Ludeña *et al.* 2013) and a decoder performs the translation task.

Moreover, several other sign to vocal language systems have been built, such as Hand in Hand (Stein *et al.* 2007 – ASL to English), MaTrEx (Morrissey 2008 – ISL to English and vice versa), SignAll (<sup>18</sup>, <sup>19</sup>, <sup>20</sup> – ASL to English). For further reading on this topic the reader is also referred to Stein *et al.* (2012), Ebling (2016) and Schmidt (2016).

In general, particular attention has been devoted in this field to the recognition task. In this regard, systems have been created, which also employ sensor gloves that are utilised by the user while signing, in order to capture and translate the performed signs. There are also translators operating from sign language that base their recognition system on the Kinect technology<sup>21</sup>.

<sup>18</sup> <https://www.signall.us>

<sup>19</sup> <https://www.youtube.com/watch?v=IE5bBLk51hw>

<sup>20</sup> <https://www.youtube.com/watch?v=hx1kn55TWpY>

<sup>21</sup> <https://www.microsoft.com/en-us/research/blog/kinect-sign-language-translator-part-1>

Furthermore, it is interesting to notice that not only there are systems able to translate from a sign to vocal language and vice versa, but systems that are able to translate from a sign language from another also exist, such as the system developed during the DICTA-SIGN project (Efthimiou *et al.* 2010).

## 2.4 Summary

The present chapter has offered an overview on Natural Language Generation and the research on sign languages in the AI field.

As far as NLG is concerned, we have seen that it is a quite recent research field, which aims at making a machine able to generate grammatically correct sentences in a natural language, and that it is grounded on choice making. On this basis, we have identified six tasks performed by most NLG systems and illustrated the main approaches to NLG.

Then, I focussed on sign languages as subject matter of AI research, introducing some generation and translation systems from and into sign languages and depicting their architectures. As for translation systems, it is important to stress that systems that translate from sign into vocal languages differ from those systems that translate from vocal into sign languages: while the former have a recognition module in order to recognise the input information in sign language, the latter have an animation module in order to produce the output in SL. It was also pointed out that early NLG and translation systems in SL field tended to be rule-based. Later on, more and more SL corpora have been created and statistical approaches have been increasingly adopted.



# Chapter 3

## The Experiment

### 3.1 Introduction

In this Chapter, the reader will be introduced to our experiment, which is based on one specific question: ‘Are we able to generate Italian from glosses of a LIS fable?’

In Section 3.2, the reader will be introduced to the domain of fables. I will describe some peculiarities of SL fables and explain how we selected the video on which this work is based. The next Section will be entirely dedicated to the chosen LIS fable, which is “The Tortoise and the Hare”. Firstly, I will briefly illustrate the content of the video and clarify how I decided to gloss it. In fact, LIS information has been organised on annotation tiers, which will be described one by one. Secondly, I will introduce the reader to the translation of the fable. The translation has been carried out on the basis of our assumptions about the text that the generator would have been able to produce. Lastly, I will evaluate the generated text and outline the main problems.

## 3.2 Fables

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, which is often not easy to process with computational systems. As far as Sign Languages are concerned, signed narratives are very common in Deaf communities and fables are popular too.

When we first decided to generate a fable in Italian from Italian Sign Language, I knew it would have been a great challenge. The reasons for this are at least threefold: Role Shift (RS), space and classifiers. Indeed, these are some peculiarities of SL narratives, including fables, which are usually not used similarly in vocal languages like Italian. For this reason, I expected them to be difficult to process in an appropriate way.

As already mentioned in Chapter 1, Deaf people tend to utilize RS very often, embodying the entity to which they are referring. Furthermore, in SL discourse the use of direct speech is widespread and the Point Of View is generally internal. It follows that signers depict the action performed by a character as they were performing it. Given this condition for the POV, the signer tends to narrate in the “here and now” with his spatio-temporal coordinates, as if narration were occurring in his or her present. This is not the case, for example, for Italian fables, which are *usually* told using past tense<sup>1</sup>.

As for the question of use of space, the reader may have observed that it is an essential feature of sign languages. For example, points of space may be used to locate entities and create anaphoric references. As observed in Chapter 1, the realisation of the sign in a particular point in space may have a specific purpose and that point may be pointed at later to refer to the same meaning or entity previously expressed by the sign. Hence, we can infer that spatial information may be extremely precise in SL. However, the precise collocation of a referent in space may not be that relevant in vocal languages. For example, in a SL dialogue between two referents it is easy to understand their spatial position. For example, we know who of the two stands on the left or on the right of the other referent, because we see it, since the signer sets the setting for us. When

---

<sup>1</sup> I wrote *usually* because Italian narratives may also be told using the so-called historical present, which allows using the present tense to narrate past events. I will not go into details on this subject in this work.

translating a SL dialogue in a vocal language, however, this information is usually not important to us: we just want to know who of the two is talking.

It is also important to stress that a single sign used for pointing (with the extended index finger) may be translated in different ways (e.g. “there”, “he”, “she”, etc.), depending on context. Thus, I considered problematic the glossing process for such signs: should I gloss them either on basis of the meaning they convey in that context or on the basis of their semantic features, as seen in Section 1.4.2?

Classifiers are signs that do not have a univocal meaning and their realisation varies depending on context. They are often used in stories in combination with RS and may make narration more thrilling, allowing interlocutors to visualise it and almost to feel like they were part of it. Furthermore, classifiers can be categorised as different parts of speech and enable the signer to convey different pieces of information simultaneously. For this reason, I pondered over how I should have annotated them, bearing in mind what information was relevant for the translation.

Moreover, let us not forget that fables in vocal languages are nowadays passed down in written form, while – as far as I know – SL fables are not. Hence, this issue should be taken into consideration too when translating, in order to choose the appropriate register.

With these considerations in mind, I started searching for the fable. We decided to choose it on the basis of two main factors: the length of the narration and the linguistic competence of the signer.

As for the former, we aimed for a story composed at least of 30 sentences, since we believe that a text that is too short cannot have a solid structure. However, for obvious reasons, we did not consider too long videos, which would have made our work far more complex and difficult.

As for the signer, we looked for a skilled Deaf signer, that would not – or as little as possible – be influenced by the Italian language.

After having taken into account different LIS videos, we opted for one of the Aesop’s fables, “The Tortoise and the Hare”.

### 3.3 “The Tortoise and the Hare”

#### 3.3.1 The video

Our LIS fable “The Tortoise and the Hare” is narrated by a skilled Deaf signer. We analysed and divided it into 45 sentences, grouped in 30 Discourse Units (DUs) that include one or more sentences.

In broad terms, as the reader may already know, the story is about a tortoise and a hare that decide to have a race. The hare starts at a very fast pace, leaves his competitor soon behind and, confident of winning, takes a nap during the race. When he wakes up, he finds out that the tortoise has almost reached the finish line. He tries to catch up, but the tortoise gets past the finish line first.

As I expected, the signer makes large use of classifiers and RS, in particular when narrating the race. The text contains dialogues, which for the most part take place between the two main characters, i.e. the tortoise and the hare. There is no explicit temporal reference, so it appears that the signer is telling the fable as it were happening in his “here and now”.

A very interesting thing to me is narrator’s position, since it seems like if he is part of the tale he narrates, as we can deduce from (1):

(1) ANIMAL WOOD ALL [...] KNOW TOMORROW MORNING EARLY  
RUN\_CL SEE+ WAIT RUN\_CL.

“All animals in the wood [...] know that tomorrow in the early morning they will run to see (the race) and wait (the beginning of the race)”.<sup>2</sup>

The use of “tomorrow” leads me to think that narrator’s spatio-temporal coordinates and those of the hare, of the tortoise and of all the animals in the wood are the same. This is not the case for fables that are written in Italian, in which narrators usually use “the

---

<sup>2</sup> The three dots in square brackets indicate text that I omitted because I consider it not relevant in (1). Text in round brackets is context information I provided to help the reader understand better the sentence. Annotations like “\_CL” and “+” will be explained in the next Section.

following day” and not “tomorrow”.<sup>3</sup> Given these considerations, I have come to reflect on the role of narrator and whether to consider him simply the person who narrates the text or the third main character of the fable who, basically, is embodied too. For the sake of simplicity in both annotation and generation phase, I chose the former option, but I personally think that further research in this field is needed.

### **3.3.2 Annotation**

I decided to annotate the fable manually, glossing signs in Italian language on different layers. Annotation has been simultaneously conducted by using ELAN and Microsoft Word. The former was used to annotate signs rapidly and in detail, synchronizing written glosses with the realisation of the corresponding signs in the video. The latter was utilised to create a file that contained glossed LIS sentences in text format. This would later be needed to create a .txt file that would have contained gloss strings, namely the input for the generator. With the benefit of hindsight, I would have simply created a .txt file and annotated directly in it. This would have made the conversion of glossed sentences to gloss strings easier.

#### **3.3.2.1 Annotation tiers**

I organised the annotation as follows. I constructed glosses on eight different 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. I will describe all of them in the next lines.

---

<sup>3</sup> Actually, I have conducted a little research on indexical and anaphoric temporal expressions in LIS recently, and some Deaf signers told me that signs corresponding to “the following day” and “the previous day” are rarely used in LIS. Such signs may be realised only when an explicit temporal reference (such as “a long time ago”) has previously been introduced in the discourse. It would be interesting to delve deeper into this matter.

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 realised to show that an action happens rapidly.

In the SYN tier I glossed syntactic NMMs, such as markers for topics and WH- and yes/no questions.

I annotated in the AGR tier specific points of the signing space used in the fable to establish thematic relations and anaphoric reference and that helped us define arguments. I also annotated in this layer points of space indicated by pointing signs.

These glosses consist of three parts: the prefix LOC that stands for “location”, a number that identifies the location sign in order of realisation in relation to the other ones and the corresponding spatial feature. Number and spatial feature are separated by an underscore “\_”. More precisely, the following spatial features described in Section 1.4.2 [+prox], [-prox], [+dist], [-dist] have been converted in this context into PROX, NPROX, DIST, NDIST respectively. An example of annotated location is LOC1\_NPROX.

As the reader may have noticed, I did not include lexical NMMs, because I thought that they would have uselessly weighed down the annotation. As we now know, lexical NMMs are particular facial expressions accompanying the realisation of specific manual signs. Since the sign would be ungrammatical if such NMMs were not produced, I decided to treat these expressions as they were incorporated covertly to the corresponding sign in the MS layer.

Furthermore, prosody head nods, fillers or hesitation markers used by the signer during the narration have not been taken into account. However, I have 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 hands and it is closely related to Action Role Shift. The attentive reader may have noticed that signs realised without the use of hands have already been introduced in Section 1.6, as Non Handed Signs (Dively 2001). I intentionally did not use the terminology used by Dively (2001) because I consider NMS and NHS to be part of two different categories. In fact, NMS seem to appear only under ARS and they seem to me to be body movements that acquire relevant linguistic meaning under RS 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, which 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 realised.

In NMS and MS tiers, I used dashes “–” when the LIS sign or information conveyed could only be translated with more than one Italian word. Classifiers are usually marked with “\_CL” at the end of the gloss. Annotations that end with “\_L” or “\_T” correspond to sign related to the hare (*lepre* in Italian) and the tortoise (*targaruga* in Italian), respectively. 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 space are introduced by the abbreviation for “index” (IX) and their annotation structure follows the annotation structure of points in 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.

I utilized the ARS tier to mark the use of the Action Role Shift related to the embodied character, such as the hare (glossed with L), the tortoise (T) and the owl (G).

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

When necessary, I used underscores in the text annotation file 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 (2):

(2) SYN \_\_\_\_\_wh  
 QRS [T\_\_\_\_\_]  
 MSRUN-LAP HOW.  
 “How is the race course like?”

The first sentence of the fable with English glosses is shown in Table 1. Translation is also shown below.

\* = LOC1\_NPROX<sup>4</sup>

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

**TABLE 1.** “English glosses of the first sentence of the fable.”

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

However, as already mentioned, I annotated using Italian words. Annotation of the whole fable is given in Appendix A.

---

<sup>4</sup> I used asterisks and wrote the locative here due to lack of space.

### **3.3.3 The target Italian version**

Before getting into the issue in question, I would like to briefly enlarge upon the notion of translation. When translating a text, there is not one correct target version only, but rather a range of possible correct alternatives, that involve, for example, deciding to use a particular term rather than another. More precisely, several stylistic choices may be made and different narrative strategies may be adopted. Let us take as an example LIS and Italian fables: if signers do not indicate temporal reference in the narrative that they are signing, the narration tends to be interpreted in present tense, as if it were occurring in their “here and now”. Italian fables are usually told in past tense instead. Hence, it is necessary to decide whether the translation should be more source-language (here LIS) oriented or target-language (here Italian) oriented.

However, let us explain in an orderly fashion how the translation process has been carried out.

The translation of the fable in its final version has not been entirely written on the first try and has gone through quite a few changes.

I simply started by watching the video several times. Then, I drew up a first draft, trying to use terms consistent with the context of fables. I did not include in my text any fillers or hesitation markers used by the signer during the narration. However, I tried to avoid a translation that was too different from the source text; for example, when possible I tried to avoid using Italian figures of speech, which are set phrases that have to be set ad hoc to be generated. Furthermore, I removed redundant words and reviewed the draft several times. The result was a well-formed Italian text.

Then, we reviewed it and adjusted it on the basis of what the generator would have actually been able to generate from glosses. Thus, our aim has been to provide an all in all correct written text that could also be produced by the generator starting from LIS glosses.

This led us to write in the text some sentences that probably strain slightly the limits of acceptability, as (3a) and (3b) on the next page show.

(3) a. la tartaruga chiede con aria perplessa : quale gara è ?  
the tortoise asks with demeanour puzzled : which race is ?  
“The tortoise asks, puzzled: «Which race is it?»”<sup>5</sup>

b. la lepre replica: " correre una gara a chi arriva prima  
the hare replies : “ to run a race to who arrives first  
? “The hare replies: « Running a speed race. »”

(3b) is odd as answer to (3a), but still quite acceptable since the communicative purpose is understandable.

There are more sentences that prove to be a little odd in Italian and sound as if they have been expressed by a non-native Italian speaker, such as (4) and (5):

(4) la tartaruga dice: " va bene accetto la gara  
the tortoise says : “ ok (1ps)accept the race  
? “The tortoise says: « Ok, I accept the race. »”

In (4), the tortoise says that he accepts the race, which it is not possible to say such a thing in Italian, from a pure grammatical point of view. He could accept *to take part in* the race or accept *a challenge* instead.

(5) La lepre venne per prima  
the hare came first  
“The hare came first”

When narrating a fable, an Italian speaker would usually use in this context the verb “arrivare”, i.e. “to arrive” and say “La lepre arrivò per prima”. However, in the video the signer uses the verb “to come”. Therefore, the best solution to us seemed to be to maintain “to come”, since we were sure that the generator would have been able to produce it.

On the whole, we decided to translate the text using past tense, where possible, and we deemed it appropriate to keep direct speech structures used by the signer in the video. We did not convert them into indirect speech also because fables are usually very rich in

---

<sup>5</sup> The first line is the generated string. The second line includes a literal translation of the string. In the third line the reader is provided with a free translation that conveys the meaning of the string.

dialogues. Furthermore, studies have shown that dialogic information is more engaging than information provided in monolog (Bowden *et al.* 2017).

The Italian translation is given in Appendix B.

### 3.3.4 The generated text

Let us now explain what our generator is able to produce and compare its output with the human-made Italian version of the fable.

Since the two versions are based on the same LIS glosses, they have substantially the same content (*what* to say), despite some omissions and repetitions in the generated text. However, they differ mainly in *how* they convey information: a native Italian speaker would immediately understand which version is the human-made one and which one is not.

When we first began working on glosses, we did know neither if we would have been able to generate some valid sentences, nor what could eventually be possible to generate. Turn taking and direct speech seemed a really tough task, not to mention the issue of grammatical tenses. The generator can instead produce a complex text with paratactic and hypotactic structures that contain dialogues between characters and it can also use different tenses properly, for the most part. Thus, we are satisfied with the generated output, which is given in Italian in Appendix C.

On the whole, the text is readable; information is conveyed in logical order and it is quite detailed compared to glosses. In fact, there are some omissions and repetitions in the text. Some Italian readers may claim that it is not that fluent, but I would say that it is a very acceptable text, given our expectations. However, the next pages are dedicated to analysing in detail the major problems.

As far as the structure is concerned, the generated text may be a little rigid, repetitive and composed of sentences that are apparently canned, as can be seen in (6) on the next page.

(6) la tartaruga risponde: " va bene la gara è adesso ?  
the tortoise replies : " ok the race is now ?  
"The tortoise replies: « Ok, does the race begin now? »"

la lepre risponde: " no domani mattina  
the hare replies : " no tomorrow morning  
"The hare replies: « No, tomorrow morning. »"

la tartaruga risponde: " sì sì va bene  
the tortoise replies : " yes yes ok  
"The tortoise replies: « Yeah, ok. »"

la tartaruga dice: " : l appuntamento è qui ? ?  
the tortoise says : " : the appointment is here ? ?  
"The tortoise says: « Are we going to meet here? »"

la lepre risponde: " sì sì qui  
the hare replies : " yes yes here  
"The hare replies: « Yes, here. »"

la tartaruga replica: " va bene grazie  
the tortoise replies : " ok thanks  
"The tortoise replies: « Ok, thanks. »"

Below our version:

(7) La tartaruga rispose: "Va bene. La gara è adesso?"  
the tortoise replied : " ok . the race is now ?  
"The tortoise replies: « Ok, does the race begin now? »"

"No, domani mattina", replicò la lepre.  
"no, tomorrow morning ", responded the hare .  
"« No, tomorrow morning », responded the hare."

"Sì, sì, va bene", disse la tartaruga.  
"yes, yes, ok ", says the tortoise .  
"« Ok », said the tortoise."

"L'appuntamento è qui?", chiese la tartaruga.  
"the appointment is here ?", asks the hare .  
"« Are we going to meet here? », asks the tortoise."

"Sì sì, qui", disse la lepre.  
"yes yes, here", said the hare .  
"« Yes, here. », said the hare."

"Perfetto, grazie", disse la tartaruga.

" perfect , thanks ", said the tortoise .

"« Perfect, thanks. », said the tortoise."

Furthermore, the tense of the introductory verbs in (6) is present, we expected simple past instead.

As for selection of other tenses, the generator usually makes the right choice. For example, (8) and (9) show correct use of imperfect and gerund, respectively:

(8) mentre **saltellava** vide improvvisamente una tranquilla  
while was-hopping saw suddenly a calm<sup>6</sup>  
tartaruga che camminava lentamente  
tortoise who was-walking slowly

"While (the hare) was hopping, he suddenly saw a tortoise who was walking calmly, slowly."

(9) la lepre si avvicina **saltellando** e chiede con un tono  
the hare moves-closer hopping and asks with a tone  
sprezzante : noi 2 possiamo fare una gara ?  
scornful : we 2 can do a race ?

"The hare moves closer (to the tortoise), hopping, and asks scornfully: « Can I race you? »"

However, in the generated version we also recognise some wrong selections of tenses, such as the imperfect tense instead of simple past. For instance, for the verb "guardare", which means "to watch", the imperfect "guardava" is generated instead of "guardò".

Moreover, some agreement errors can be noticed. For example, in (10) the hare is illustrating to the tortoise the route of the race, explaining what they are both going to do. The first verb in the first sentence is "cominciamo", 1pp of the verb "cominciare" (to begin). Logically, the other finite verbs should be inflected to the 1pp, yet this is not the case since they are inflected to the 3ps<sup>7</sup>:

(10) poi cominciamo a correre e **fa** il giro  
then (1pp)start to run and (3ps)does the lap  
"Then we start and (3ps) runs the lap"

---

<sup>6</sup> As the reader may have inferred, Italian is a pro-drop language.

<sup>7</sup> "1pp" stands for "first person plural", "1ps" stands for "first person singular" and so on.

poi **continua** a correre fino a la vecchia casa abbandonata  
 then (3ps)continues to run to the old house abandoned  
 “Then (3ps) continues to run and (3ps) runs to the old abandoned house”

quindi **finisce** il giro qua  
 then (3ps)finishes the lap here  
 “Then (3ps) stops here.”

As already indicated, the generator can generate direct speech correctly. However, some errors occur, as we can see in (11), where the tortoise:

(11) chiede : rinuncia per paura ?  
 asks : gives up for fear ?  
 “(3ps) asks: Did he give up because he is scared?”

risponde: " no dice a la gara ci partecipava  
 replies : “ no says to the race to it (3ps)took part (imperfect)  
 \* “(3ps) replies: « No (3ps) says (3ps) was taking part in the race. »

Our version:

(12) "Io rinunciare per paura? No. Io alla gara ci partecipo".  
 “ I to give up for fear ? No. I to the race to it take part ”.  
 “Me, giving up because I’m scared? No, I’m taking part in the race.”

So, as we can see, sentences in (11) are not generated as part of the same discourse turn. Furthermore, the second sentence is ungrammatical and the word “dice” (3ps of the introductory verb “to say”) is uselessly inserted. In addition, verbs are inflected to the 3ps; they should be inflected to the 1ps instead.

Moreover, anaphora resolution is not always an easy task and agreement in gender may not be marked correctly. This is the case with the nouns “hare” and “tortoise”, which in Italian are both feminine and, when combined, result in a feminine plural form. At some point in the generated fable, the two animals converse and then end their conversation. Right after, follows (13a) with the masculine plural article “i” and not (13b) with the feminine “le” as used in the human written translation:

- (13) a. **i** due si affiancarono  
 b. **le** due si affiancarono  
 the two come side by side  
 “The two (of them) come side by side”

The same error recurs a few lines below. In fact, an owl comes out of the blue and asks them if they are ready (to begin the race):

- (14) ora arriva un gufo e dice : voi due siete pronti ?  
 now comes an owl and says : you two are ready ?  
 “Now comes an owl and says: « Are you two ready? »”

As we can see in (14), the generator produces by default the masculine suffix “-i”, resulting in the generation of “pronti”, and does not produce the feminine suffix “-e”, that would have created the correct form “pronte”. However, this is due to the fact that genre specifications have not been added into the LIS glosses, since LIS signers do not use gender to refer to interlocutors present in the context.

In addition, some adjectives are generated as predicative adjectives (15a), in a context where their use with attributive function (15b) would be more appropriate:

- (15) a. sorpassò l albero grosso  
 b. sorpassò il grosso albero  
 “(3ps) passed the big tree”

There are also some orthographic inaccuracies that involves the missing combination of preposition and article, such as the feminine “a”+“la” that should result in “alla” and the masculine “a”+“il” that should result in “al”.

The question of definiteness may be a problem too, as we can see in (16):

- (16) nel bosco viveva una lepre **la** lepre altezzosa  
 in the wood lived a hare the hare haughty  
 “In the wood lived a hare, the haughty hare.”

If a noun has already been introduced in the discourse, the generator associates it with a definite article. However, this should not be always the case. For example in (16),

which is the first sentence of the fable, the apposition “la lepre altezzosa” should be replaced with “una lepre altezzosa”.

### **3.4 Summary**

I started my thesis with a question in mind: Are we able to generate Italian from LIS glosses? The present Chapter was devoted to finding it out and, at this point, we can claim that we are. The first part of this Chapter has given an overview of some characteristic of SL fables. Then, I have introduced the reader to the fable that we chose for this work, i.e. a LIS video of “The tortoise and the hare”. After having explained how I have annotated it, I presented the Italian translated version of the glosses, which was also used as target translation. This translation has gone through quite a few changes. In fact, we adjusted it on the basis of what we expected the generator would have been able to produce. The last Section of the Chapter was dedicated to describing and evaluating the generated text. All in all, we can claim that we are satisfied with the generated output, although there are issues that need to be solved.

# Chapter 4

## The GENLIS generator

### 4.1 Introduction

In the previous Chapter the reader was introduced to our experiment on generation.

The last Chapter of this thesis is completely devoted to our GENLIS generator, a machine that generates Italian from LIS glosses.

In the next Section we describe the generation mechanism. More precisely, we introduce glosses and semantic forms in Prolog, which are received as an input for the generation. Furthermore, we present the most important algorithms that are integral part of the generator and we focus on mapping verb tense and mood.

In Section 4.3, I explain the main problems encountered in the generation and I propose some possible solutions.

## 4.2 Generation mechanism

A general overview of the generator mechanism is shown in Figure 1:

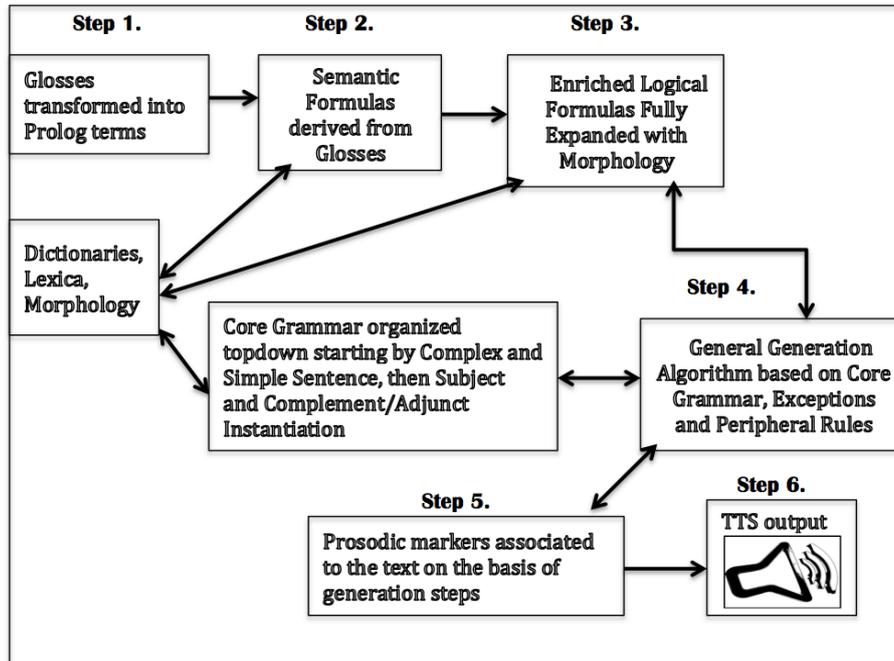


Figure 1. The generation mechanism in GENLIS

Figure 2 illustrates Step 3. and Step 4. more in detail:

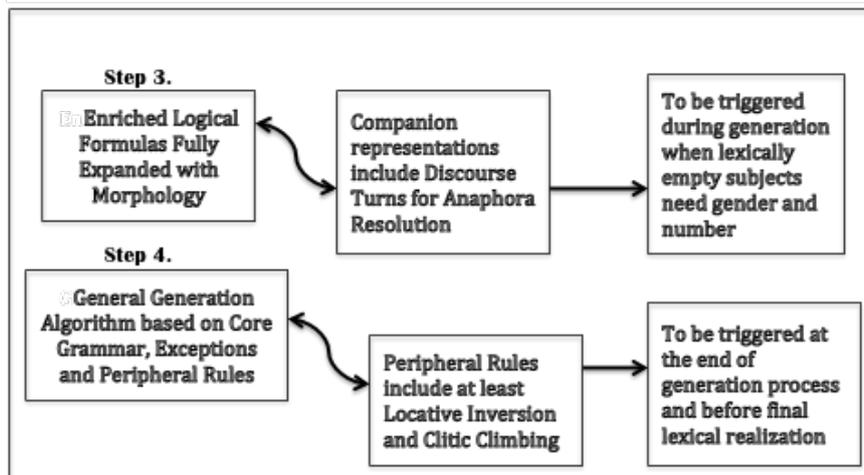


Figure 2. Detailed overview of Step 3. and Step 4.

As can be seen in Figure 1, a Text To Speech (TTS) system is included in our project, but not available yet. Therefore, we will not deal with this topic in this thesis.

The following pages are dedicated to some key points of the generation mechanism.

#### 4.2.1 Manual glosses and semantic forms in Prolog

GENLIS is written in the logic programming language Prolog. Obviously, this is not the case for manual glosses, which as we have seen are basically multi-layer text annotations written in tables.

In order for the multi-layer glosses to be analysed by the generator, it has been necessary to transform them into one-layer strings. Thus, each annotation tier has been inserted in slots in a string, as follows:

```
gls (DUInd, Aff, Adv, Syn, Agr, Nms, Ms, Ars, Qrs)
```

where the functor is `gls`, abbreviation for “gloss”, and contains a list composed of different terms. `DUInd` indicates the Discourse Unit. The slots `Aff`, `Adv`, `Syn` contain annotated information about affective, adverbial and syntactic NMMs. `Agr` identifies location and agreement of signs. `Nms` and `Ms` contain Non-Manual Signs and Manual Signs respectively and are expressed in a tokenized sequence between apostrophes ‘ ’ as atomic objects. `Ars` and `Qrs` identify the occurrence of Action Role Shift and Quotation Role Shift.

However, generation takes place in semantic forms, which are composed by the main predicate, prepositional attributes (such as mood, negation, verbal tense, etc.), arguments and adjuncts. Furthermore, each argument has its own internal structure.

Semantic forms constitute the string that is fed as an input to the generator and then processed, in order to generate Italian sentences.

We will describe below both the process of transformation of glosses into semantic forms and the structure of semantic forms.

First of all, we elaborated conversion rules from manual glosses to semantic forms that would be needed to generate sentences in natural language. We avoided indicating specific features that would make the forms difficult to read and understand. More precisely, we did not indicate tense, mood and diathesis of verbs, number and genre of nouns and semantic role of oblique arguments for the generation of prepositions.

We decided conventionally to generate sentences with active diathesis, in past tense and indicative mood. However, there are several factors to take into consideration: direct speech and questions, for example, are always expressed in present indicative. Furthermore, morphological features of nouns are always singular, unless otherwise indicated in glosses, and genre is derived from lexical genre. Past verb tense is derived on the basis of the aspect of the lexical verb; in particular, state and action verbs are expressed in imperfect tense and the other verbs in past tense (“passato remoto” in Italian).

Semantic forms are structured as Prolog terms. Consistent with First-Order Logic (FOL), each term represents the content of a semantic proposition and is preceded by the functor PROP<sup>1</sup>. PROP is the abbreviation for “proposition” and contains a defined number of slots that mark semantic and pragmatic components included in glosses. More precisely, in the first slot we may find pragmatic components like interjections of surprise or other affective and emotional aspects, intrasentential elements, discourse markers and adverbs with scope on the verb or on the entire sentence.

Let us now focus on the arguments structure. With the exception of SUBJECT and OBJECT, arguments are introduced by a functional marker, such as OBL for oblique arguments, FCOMP for sentential complement, VCOMP for verb complement and XCOMP for predicative complement. Moreover, argumental heads may contain modifiers, which are introduced by the marker MOD, or specifiers, which are usually indicated in brackets. If the argument is an expression of the affirmative or the negative, the expression becomes the only term of the argument list.

Moreover, direct speech may be deprived of any introductory verb, which needs to be generated in Italian instead and may assume different meanings depending on context, as we will see on the next pages.

---

<sup>1</sup> In this Chapter, some words are capitalized for stylistic purposes. They are not intended to be read as LIS glosses as in the previous chapters, unless otherwise indicated.

Conversion rules from manual glosses are shown below:

0. Identify elements that modify the main predicate, adverbs or discourse markers.
1. Insert the first verb you find.
2. Retrieve lexical verb aspect and create mood/time matrix.
3. The verb may be preceded by a location, which may be marked by a specific deictic term on the basis of type.
4. Speech act may vary:
  - PRESENTATION = WHO?
  - DIRSPEECH for direct speech
  - QUESTION if the sentence is a question
  - PERLOCutive when the owl starts the race by saying “3, 2, 1... via!”
  - ITERATive for repeated forms of the same verb
  - the default STATEMENT.
5. Insert arguments into a list [ ]:
  - 5.1 The subject in the first slot may be unexpressed. If so, it is marked with “little\_pro”: morphological features are retrieved from the subject in previous sentence.
  - 5.2 In case of direct speech, arguments may be interjections or statements/negations.
  - 5.3 The object may be a complement sentence marked FCOMP, an interrogative complement sentence marked QCOMP or an infinitive sentence marked VCOMP.
6. Oblique arguments or adjuncts are marked OBL and may contain in their first position either a preposition, if expressed overtly in manual glosses, or a semantic marker, and the lexical head in their second position.
7. Nouns may have specifiers, such as “gara-[quale]” (which means “race-[which]”) and modifiers, which are marked MOD.
8. Adverbs such as locative deictic adverbs are marked AVV.
  - 8.1 Gerundives are marked AVV too and contain the corresponding infinitive verb.
9. PROpositions may be coordinated (COORD) or appear in sequence without markers (IPOTAS). These tags are inserted first, before the PROP tag.

As an example, the semantic form of the first sentence is shown below:

```
lepta(1,1,prop(nil,vivere,attivita,present,
  [lepre, lepre-[mod-[altezzoso]], obl-[luogo,bosco]])).
```

## 4.2.2 The Generation Algorithm

The generation algorithm is basically a “core” grammar algorithm with peripheral rules for exceptions. The generation is organised around the following steps:

### *First step: simple vs. complex sentences*

A subdivision at utterance level between simple sentence and complex sentence is created. Complex sentence input includes coordinate sentences and subordinate sentences, which can be both in turn simple sentences with a discourse marker at the beginning.

### *Second step: organising sentences*

Simple sentences are then organised into different types, based on diathesis which may be active and passive, but also based on speech act types, a separate call for questions or exclamations, which may require a specific sentence organization.

### *Third step: generating the subject*

Each simple sentence generates the subject and then calls the rest of the sentence; in case sentence type is question, a check is made to see if the subject is constituted by a WH- word; empty subject pronouns are classified as “little\_pro”.

### *Fourth step: nominal expressions*

All nominal expressions - both SUBJECT and OBJECT and OBLIQUES – can be modified by simple modifiers, multiple modifiers, and relative clauses. All of them are structurally attached to the nominal head because they are semantically and morphologically dependent on the head. In fact, adjectivals require feature agreement, which needs to be restricted before generation in order to prevent failures. As to relative clauses, their internal arguments may require the same type of information, in particular in case the argument controlled by the relative pronoun – which may be unexpressed –

is the SUBJECT. Relative clauses may also be governed by an adjunct relation, but this is not the case in our story. In order to realise the appropriate word forms, the morphological features of the nominal head governing the relative clause are passed to the clause level as BINDER bundle of features, which may be used by the Verb Complex and realized as SUBJECT or OBJECT features.

*Fifth step: choosing verb type*

The remaining generation component is governed by the choice of verb type, which has three possible options: copulative, transitive and intransitive. The verb complex receives semantic and morphological information from the subject if present, be it a nominal or pronominal head, or simply an empty subject which however may have morphological features, person, number and possibly gender.

Choosing the correct verbal complement structure may be dependent on subject semantic categories, which are also passed to the verbal complex. Semantic features are checked by matching subcategorisation information stored in the lexicon for each possible structural outcome. For instance, a verb like *dire* has the following entries in our lexicon:

```

Gramm_cat=trans, Semantic_cat=riport_dir,
FIRST = [sn/sogg/actor/[umano, informa, astratto],
         vcomp/prop/di/[sogg=sogg/actor]],
SECOND = [sn/sogg/actor/[umano, informa, astratto],
          sn/ogg/tema_nonaff/[informa, astratto]]
THIRD = [sn/sogg/actor/[umano, informa, astratto],
         sp/ogg2/esperiente/a/[umano],
         f/fcomp/prop/[sogg=sogg/agente, sogg=x]]
FOURTH = [sn/sogg/actor/[umano, informa, astratto],
          f/fcomp/propq/[sogg=sogg/agente, sogg=x]]

```

As can be noticed, there are four separate complement structures:

- i. vcomp = INFINITIVAL
- ii. ogg = DIRECT\_OBJECT
- iii. ogg2 = INDIRECT\_OBJECT (dative) + f/fcomp = SENTENTIAL\_OBJECT
- iv. f/fcomp = SENTENTIAL\_OBJECT

which are however characterised by the same grammatical category, TRANSitive, and the same conceptual and semantic category, *riport\_dir*, that is a reporting verb that can be used also for direct speech introduction.

This also applies to other verbs that may undergo “intransitivisation” like *mangiare*, but also to verbs with different complement structures but identical categorisations, like *considerare* and *dipingere*:

```
FIRST= considerare, trans, soggettivo,
      sn-sogg, esperiente, [umano],
      sn-ogg, [nn], tema_bound, [_],
      ncomp-[sogg=ogg], [nn], prop, [_]
```

```
SECOND=considerare, trans, soggettivo,
      sn-sogg, esperiente, [umano],
      sn-ogg, [_], tema_bound, [_],
      xcomp, [_], prop, [_]
```

```
FIRST= dipingere, trans, attivita,
      sn-sogg, agente, [umano],
      sn-ogg, [_], tema_aff, [oggetto]
```

```
SECOND=dipingere, trans, risultato,
      sn-sogg, agente, [umano],
      sn-ogg, [_], tema_aff, [oggetto],
      acomp, [_], prop, [_]
```

where we see for *considerare* open complements like NCOMP (a nominal predicative complement) or XCOMP (a label for generic open complements including infinitivals); and (no) open complement controlled by the OBJECT nominal for *dipingere*. All open complements require morphological features to match, so that the call for complement structures has to impose agreement for those features.

This can be different for other verbs where grammatical category may vary, as is the case for *accennare* that goes from intransitive to transitive:

```
pred_vs(accennare, intr, attivita, sn-sogg, actor,
[umano], sp-ogg2, a, tema_nonaff, [oggetto], nn, [nn], nn,
nn).
pred_vs(accennare, trans, attivita, sn-sogg, actor,
[umano], sp-ogg2, a, tema_nonaff, [umano], f, [nn], prop,
[sogg=sogg/actor, sogg=x]).
```

or for a verb like *apparire* that goes from copulative to unaccusative or *scappare* from unaccusative to intransitive:

```
FIRST=  apparire, cop, stato,
        sn-sogg, actor, [animato, umano],
        ncomp, [_], prop, [_],
SECOND=apparire, inac, risultato,
        sn-sogg, actor, [evento, animato, umano]
```

```
FIRST=  scappare, inac, achiev,
        sn-sogg, tema_aff, [animato, umano]
SECOND=  scappare, intr, attivita, sn-sogg, agente, [animato,
        umano], sp-obl, da, malef, [animato, umano,
        ferocious, evento, attivita, luogo]
```

The call for verbal complex and complement structures will be different in these cases, so that when none of the expected complement is met by the generator a failure will ensue and another call will be attempted.

Our lexicon is organised around a limited number of entries, around 1000 for most frequent lexical entries according frequency dictionaries, and another extended set of manually annotated entries, around 9000 for the remaining less frequent but always non rare entries, which have a different feature and argument organization. Here is the set of lexical entries for *abbandonare*, reported as the verb root *abbandon*, followed by the paradigm class – 1<sup>st</sup> conjugation –, a label for the now extended grammatical class that includes “inherent reflexive”, and an extended set of semantic and conceptual classes.

```
FIRST=  abandon,1, refl_in, activ, hyper,
        [np/subj1/agent/[rifl,+hum],
        pp/obj2/goal/a/[+abst]]) .
SECOND=abandon,1, refl_in, activ, hyper,
        [np/subj1/agent/[rifl,+hum],
        pp/obl/locat/su/[-ani,-abst]]
THIRD=  abandon,1, tr, accomp, posit,
        [np/subj1/theme_unaff/[-hum,+abst],
        np/obj1/theme_aff/[+ani,+hum]]
```

FOURTH=abbandon,1,tr,achiev,posit,  
 [np/subj1/agent/[+hum,+ani],  
 np/obj1/theme\_unaff/[-ani],  
 pp/obl/locat/su/[-ani,-abst]]

FIFTH= abbandon,1,tr,achiev,posit,  
 [np/subj1/agent/[+hum,+ani],  
 np/obj1/theme\_unaff/[+ani,+hum],  
 pp/obl/locat/a/[]]

As can be seen, we can have a two place argument structure, but also a three place one with obliques; a semantic classification for HYPER, and another one for POSIT (positional) involving locatives; Accomplishments, Achievements and Activities. In particular Aspectual categories are very important – as said above – in the choice of verbal morphology regarding Tense and Mood; while semantic and conceptual class may also be relevant in case a sentential complement is present, as will be clarified below.

*Sixth step: generating the verbal complex together with its complements and adjuncts*

Generating the verbal complex requires precise morphological information as to the Tense and Mood to be realised. In particular, simple vs. composite verbal complex may be realised, which in turn require specification of the appropriate auxiliary verb: *essere* for passive, reflexive, inherent reflexive and unaccusative classes, *avere* for active transitive and intransitive classes. Morphological information from the SUBJECT is also required in case of auxiliary *essere* in order to generate the appropriate past participle. The same is required from the OBJECT in case of pronominalization processes of the nominal head into a clitic pronoun, which however requires decisions that can only be made by a full-fledged pronoun resolution system.

As to Person, this may be available in case the SUBJECT is lexically expressed. Empty pronouns on the contrary do not realise Person feature, which is by default set to 3<sup>rd</sup>. Special cases are constituted by Imperative mood and Direct Speech. Imperative mood requires 2<sup>nd</sup> person to be realised if the command or instruction is addressed directly to the interlocutor. But there are commands in the fable addressed by the owl to both competitors, the hare and the tortoise, to start the race. In this second case, 2<sup>nd</sup> person plural is required. However, 1<sup>st</sup> person plural is also acceptable. Introducing 2<sup>nd</sup> person

is not an easy task and we haven't been able yet to find a linguistically motivated trigger to do it.

When the verb is finally selected and lexically realised, it is checked for agreement with SUBJect morphological features. This may cause failures until the appropriate verb form is produced.

The following call is the one for complements and adjuncts.

The input guides the selection according to its shape: nominal and sentential complements are made up of a four or five slots list, while an oblique may be constituted by a list containing five or six slots; a simple modifier has only two or three slots. Finally adverbials or interjections consist of one or two slots but contain a special label as unique identifiers.

Sentential complements may be simple sentences preceded by a complementizer, which is locally generated; or they may be direct questions. In this second case, a question mark is added at the end. The two complement types are marked by a special label identifier FCOMP and QCOMP.

A special case may be constituted by WH- questions as sentential complements, requiring a local WH- expression to be generated before the verb also in case it is an adjunct – i.e. *when, how, where*. These pronouns would be positioned after the verb in the logical form built from semantic forms. So they need to be raised, i.e. removed from the complement structure and generated in the appropriate position.

#### *Peripheral and special rules*

Peripheral rules are required to generate special stylistically marked structures, like Subject Locative Inversion with presentative structures, and complements realised as clitics, which need to be positioned before the verb. In both cases we implemented the rules to act at the end of the generation process. The SUBJect – Locative is used for the first sentence of the fable, when the hare is presented and appears on the scene as living in the woods. This is a typical introductory sentence for many fables or children stories and has all the required linguistic features: the protagonist is unknown and is realised as an indefinite nominal structure; the verb is unaccusative or intransitive, in this case *vivere* used intransitively; the sentence is completed by presence of a Locative adjunct,

“*nel bosco*”. The main linguistic elements are all generated in their base structure, they are identified and displaced in order to produce a presentation structure where the Locative comes in first position followed by the verb complex and then comes the subject nominal and finally the rest of the sentence, which in this case is an apposition.

The second case of peripheral rule is the one involving the clitic pronoun “*ci*” for a locative or a dative repeated in the same complement structure, and the governing verb *partecipare*. The clitic is generated after the verb and then it is scrambled before it.

Structures that require special rules to be implemented include so-called Open Complements and Open Adjuncts. Open Complements are predicative complements of copulative verbs, as in “*siete pronti*”; Open Adjuncts are state adjectives like *tranquillo*, which require gender/number agreement with the SUBJECT as in “*la tartaruga guardava tranquilla*”. Both cases require SUBJECT morphological features to be visible in the Complement/Adjunct section of the generator in order to select or restrict the appropriate word form.

### 4.2.3 The Algorithm for Definiteness Assignment

In order for the generation to work properly, the feature “definite”, “indefinite” or “zero” must be decided automatically and inserted in the list of features associated to each nominal expression, be it the primary head as with subjects and object, be it secondary with obliques where the noun phrase is governed by a preposition. The list of features includes morphological, semantic and informational features as follows:

[Def, Spec, Num, Head]

Def contains the information about definiteness if the head is a noun, otherwise it is substituted by TOP in case the head is a pronoun. Spec contains information on quantification and any linguistic element that may be expressed by a quantifier. Num is associated to the morphological feature of Number.

The Algorithm for Definiteness Assignment (ADA) is based on two parameters: the type of constituent and the semantics associated with the noun. The semantics is taken

from a set of different sources due to their dimensions, which are insufficient to cover all nominal expression of the fable. We have been using the lexical-semantic database ItalWordNet<sup>2</sup>, the list of words annotated by semantic Supertags and lexical categories. The code is shown below:

```

known_def(Func,Head,def):-
    def(Head,_),
    !.
known_def(Func,Head,Def):-
    except(Head,Def),
    !.
known_def(obl,Head,def0):-
    ord(Head,L,F,N)
    ;
    nclp(Head,Pol,Feats),
    member(part,Feats),member(plac,Feats)
    ;
    nclp(Head,Pol,Feats),
    member(expr,Feats),member(mnt,Feats),
    !.
known_def(Func,Head,Def):-
    dict_feat(Head,Def),
    !.
known_def(Func,Head,ndef):-
    asserta(def(Head,ndef)),
    !.

except(referente_1,pro).
except(referente_2,pro).
except(referenti_2,pro).
except('3 2 1 ... via',def0).

dict_feat(Head,def):-
    ord(Head,L,F,N)
    ;
    nclp(Head,Pol,Feats),
    member(expr,Feats),member(mnt,Feats)
    ;

```

---

<sup>2</sup> <https://www.cnr.it/it/banche-dati-istituti/banca-dati/442/italwordnet-iwn>

```

nclp(Head, Pol, Feats),
member(part, Feats), member(plac, Feats)
;
nclp(Head, Pol, Feats),
member(part, Feats), member(liv, Feats), member(fnct, Feats)
;
findall(F, sst(Head, noun, F), Fs),
(
member(location, Fs), member(group, Fs);
member(body, Fs);
member(time, Fs)
),
!.

```

The main call is `known_def`, which is used to assert in memory the type of definiteness associated to a nominal head. When a noun is met for the first time it is asserted as NDEF i.e. “indefinite”, unless it belongs to a set of exceptions and special semantic classes, which in our fable include the following words:

*appuntamento, vergogna, orecchio, giro, sinistra, destra, primo, tono*

All these nouns may be associated with Definite or Zero Definite according to semantic features and in some cases to special conditions. In addition, expressions like “referente\_N” where there is a number varying from 1 to 2, are regarded pronouns. Frozen expressions like “3 2 1 ... via” are marked with definiteness zero. The number belonging to the class of ordinals is tagged with zero definiteness only in case they are included in an oblique governed by *arrivare*. The same choice of zero definiteness applies to nominal expressions characterized by an abstract feature, which in ItalWordNet is represented by MNT (= “mental”) and EXPR (= “expressive”) tags. It also applies to words indicating location tagged by PART (= “part”) and PLAC (= “place”).

Another interesting class is constituted by words belonging to *Body\_Part* like *orecchio*, which are tagged as definite and characterized by features PART, LIV (= “living”) and FNCT (= “function”); the same applies to nouns belonging to *time* semantic class, like days, months, but also *appuntamento* whenever they are included in a nominal

constituent. Of course, all adverbial like expressions and interjections are not considered and do not receive a list of morphological and semantic tags as said above.

#### 4.2.4 The Algorithm for Narrative direct speech speaking verb type

In the glosses, direct speech is introduced always by the same verb *dire/say*. It may also be deprived of any introductory verb, which in our case needs to take into account the semantic content of the current utterance. In addition, depending on current discourse turn speaker, this verb may assume different meanings, which are strictly discourse related. So either *dire* is substituted by a contextually determined verb or a verb is introduced which was not present. These verbs belong to Answering semantic type and are: *rispondere* or *replicare*, in case the speaker is answering a question from previous discourse turn; otherwise the predicate may belong to the Asking type, *chiedere/domandare*, in case the current turn is made of a question; eventually it may also be *dire* in case of the previous turn was a yes/no question, or the current turn is a statement. Finally with exclamations it may be *esclamare*.

The algorithm is part of the “convert” file, the conversion algorithm that starting from glosses organizes them into semantic forms. It is activated after all conversions have been already made by the recursive call,

```
substitutespeechact (Output, Revturns, NewTurns)
```

The real call is the following one, where semantic forms are each one accompanied by a vector representation of each turn, with a current topic speaker, a speech act associated to the current utterance, and a main predicate, then a Discourse Unite index, a Sentence index and a proposition index, like this: Head-Spac-Pred-Du-Sn-N. These representations have been asserted into memory in a Prolog database and may be extracted easily.

```
subspac (LogForms, Turn, DefTurn, Old)
```

which receives Semantic Forms and checks to verify whether the current verb is *dire*:

```

subspac (Logs, Turn, NewLog, Old) :-
    Turn=Head-Spac-Pred-Du-Sn-N,
    Pred=dire,
    extrlog (Logs, Du, N, NewLog, Old)

```

Here `extrlog` makes a set of new checks and then calls the final predicate, which is now headed by `Du`, i.e. the index of current Discourse Unit:

```

modifypred (Du, Pred, Args, Body, NewBody) ,

```

It also contains the current governing predicate, the arguments of current predicate in the `Body` variable, the Arguments of the first sentential complement (if any) of the `Body` variable in the variable `Args`, and finally the variable `NewBody` that will contain the modified version of the arguments. The first call to `modifypred` checks to see what is the speech act of the first proposition chosen:

```

is_list (Body) ,
member (Arg, Args) ,
Arg=[quest|_],
Body=[Dirs, T, M, Pred, Args] ,
NBody=[Dirs, T, M, chiedere, Args] ,

```

In this case the `Pred` is substituted by a predicate of the Asking type, *chiedere/domandare*. The second call is the most important one and is accompanied by a check of the previous turns like this:

```

is_list (Body) ,
member (Arg, Args) ,
    checktopicspac (Du, Spac) ,
Body=[Dirs, T, M, Pred, Args] ,
    (Spac=quest, Pred1=rispondere; Pred1=replicare) ,
NBody=[Dirs, T, M, Pred1, Args] ,

```

The call to verify previous turns is `checktopicspac`, which is used to look into the database of turns:

```

checktopicspac (Du, Spac) :-
    findall (Top-Spac, turn (Top, Spac, Pred, Du, _, _), Tops),
    member (Top-Spac, Tops),
    Spac=quest,

```

The search is interrupted in case the current utterance contains a question as one of its sentential complements. Then the second call searches the turns database:

```

checktopicspac (Du, Spac) :-
    turn (Top0, Spac0, Pred0, Du, _, _-_),
    Du0 is Du - 1,
    findall (Top-Spac, turn (Top, Spac, Pred, Du0, _, _), Tops),
    member (Top-Spac, Tops),
    Top \= Top0,
    Spac=quest,

```

At first it extracts the previous turn and then it checks to see whether the current topic is different from the one asserted in the previous turn; finally it checks whether the speech act is a question. In this case the main predicate is modified into one of the Asking type.

#### **4.2.5 Mapping Tense and Mood from Speech Act and Verbal Aspectual Lexical Properties**

Every fully expressed proposition has a verb that needs semantic and morphological features. While Person, Number and Gender may be inherited from the Subject, Tense and Mood are semantically and pragmatically determined. We have used lexical properties and discourse related (pragmatic) properties to assign Tense and Mood together with general consideration defined on the basis of narratological criteria. A fable or children story may be expressed using Indicative Present or Past tense (or “passato remoto”), however contextual conditions may impose constraints that require other Mood and Tense to be assigned. We may need to use Future tense, Imperative mood, Past tense (or “passato remoto”) rather than Present tense. A first subdivision of

Mood-Tense assignment depending on Speech Act is shown below, a second subdivision follows according to Lexical Aspectual properties.

**1. Presentative constructions**

choosemoodtense (attivita, present, imp, ind) .  
choosemoodtense (achievement, present, pres, ind) .

**2. Perlocutive utterances**

choosemoodtense (process, perloc, pres, imperat) .  
choosemoodtense (attivita, perloc, pres, imperat) .

**3. Question + Exclamation**

choosemoodtense (achievement, exclam, pres, ind) .  
choosemoodtense (stativo, question, pres, ind) .  
choosemoodtense (accompl, question, fut, ind) .  
choosemoodtense (achievement, question, pres, ind) .  
choosemoodtense (modal, question, pres, ind) .

**4. Illocutive constructions**

choosemoodtense (achievement, illocut, fut, ind) .

**5. Direct Speech constructions**

choosemoodtense (achievement, dirspeech, pres, ind) .

**6. Statements**

choosemoodtense (process, statement, pres, ind) .  
choosemoodtense (stativo, statement, pres, ind) .  
choosemoodtense (accompl, statement, pres, ind) .  
choosemoodtense (achievement, statement, pass\_rem, ind) .  
choosemoodtense (attivita, statement, imp, ind) .

We distinguish Perlocutive from Illocutive verbs on the basis of the pragmatic nature of the action expressed: instructions on how to carry out a task are tagged Perlocutive and are enacted with Imperative mood. Illocutive expressions are tagged when the utterance expresses a decision or a wish to come true and are placed in the future Tense. Then, as a general rule, Activities are realized with Indicative “Imperfetto”, while Achievement use Past\_tense (passato\_remoto). The remaining cases are all realised with Indicative Present.

### 4.3 Main problems

During the creation of the above mentioned algorithms we encountered several problems. These issues are related to the generation of the Italian output text starting from LIS glosses. We bypassed most of them by modifying information in multi-level manual glosses and reviewing the target translation. Obviously, modifications have been made within reason. In fact, there is not only one correct translation and a text may be translated in different ways. Therefore, we had a certain margin that allowed us to use words that would have been more easily producible by the generator, and still convey the same meaning and maintain the communicative goal of the signer.

However, it is worth to point out that we did not *translate* a fable text from LIS glosses to Italian, but rather *generate* it. I do not utilize the term “translate” for two main reasons. From a linguistic point of view, we did not actually translate from LIS, since we did not use a visual recognition system that allowed translating the visual input. From a computational point of view, we did not use RNNs<sup>3</sup> and other deep learning applications.

As we now know, LIS is a real natural language, which uses signs and not words. As the other sign languages, LIS utilizes the visual-manual modality and does not have a standardized and recognized writing system. Thus, I knew that glosses and their organization would have been essential in order to generate an acceptable Italian text.

Let us now illustrate some elements that had to be processed ad hoc, since we have not been able to generate them in any different way.

Manual glosses are a very subjective writing system: 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.

My aim was to avoid creating a too restrictive and informative MS gloss that conveyed information that could be conveyed in other layers. Therefore, I organised gloss tiers trying to distribute different types of information on different levels. In this way, I

---

<sup>3</sup> Recurrent Neural Networks (RNNs) are specific artificial neural networks in which the output of the previous step is the input of the following step. They can be used in predictive analysis.

would have limited ad hoc procedures. In some cases, however, this was not possible. Let us take a practical example.

At the beginning of the fable, the signer tells us that the hare is hopping, by embodying him. Then he produces the sign in Figure 3 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 realisation of the sign and the picture right is its end.

Figure 3 is a classifier: in this context, the signer’s hands represents the eyes of the hare, which move right, toward the position in which will be then located the tortoise. During the realisation of this sign the movement of the hands is rapid and sharp and represents the movement performed by the eyes of the animal. Clearly, the conveyed information 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, I glossed the sign in Figure 1 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 we decided to translate the information as follows: “While the hare was hopping, he suddenly saw a tortoise”.

Since the adverbial information was not extracted from the MS layer in glosses, it could not be processed separately: there was no overt correspondence between the occurrence in my translation and the corresponding signed information in LIS glosses, therefore we had to process it 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 that have very specific functions and that have then to be transformed in a sequential string. There is 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 organisation of glosses and expand the MS layer, dividing it into the four levels that correspond to the formational parameters introduced in Chapter 1: HandShape, MOVement, LOCation and Palm Orientation. When necessary, modifications to the standard realisation of the sign may be indicated in these tiers, as shown in Table 1:

<b>MS</b>	EYES
HS	F
MOV	MOVE RAPIDLY
LOC	
PO	TO THE RIGHT

**Table 1.** “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 1, for example, depicts a classifier, i.e. a non-standard sign, which is realised with a different handshape compared to the handshape used in the realisation of the generic sign for “eyes”.

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

LOC remains empty in this case, but could be used for agreement purposes and anaphora reference, substituting for the AGR layer I created in my glosses.

Finally, PO contains the meaning conveyed by changes in palm orientation, “to the right” in Table 1.

In this way, the generation of all the information carried by the classifier should have been 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 organise information conveyed by other classifiers or signs too. Perhaps it is too restrictive or unsuitable.

There are other signs involved in ad hoc procedures 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, I consider the sentence in (1) a little odd if used to narrate a fable:

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

As Italian speaker, I prefer the use of “arrivò” (3ps “arrived”) instead. On the contrary, (2) in LIS is totally acceptable:

\_\_\_\_\_wh  
 (2) POI VENIRE CHI. GUFO.  
 then to-come who? Owl .  
 “Then comes an owl.”

It is also important to notice that there is actually not only one gloss tier that conveys consistently 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 2, written in English glosses:

\* = LOC2\_NPROX

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

**Table 2.** “The hare pants and looks right”

In Table 2, the most part of relevant information is conveyed through NMMs. In fact, PAWS(L)\_CL is a Noun and Limb/Body Part Classifier, which helps to indicate that the hare is being embodied. However, the layers that explain what the animal does (Figure 4) are actually NMS and AGR.



**Figure 4.** Embodiment of the hare panting and looking right

The same applies to Table 3 and Figure 5:

\* = GETS-READY-TO-START

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

**Table 3.** “The tortoise gets ready to start”



**Figure 5.** Embodiment of the hare getting ready to start the race.

Therefore, there should not be an overt generated correspondence of this kind of manual classifiers in the output text. Our generator produced PAWS(L)\_CL and PAWS(T)\_CL as verbs instead, and did not take into account the “gets ready to start” part.

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

As far as pragmatics is concerned, it is important to remember 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.

I annotated this sign in LIS as “to walk”.



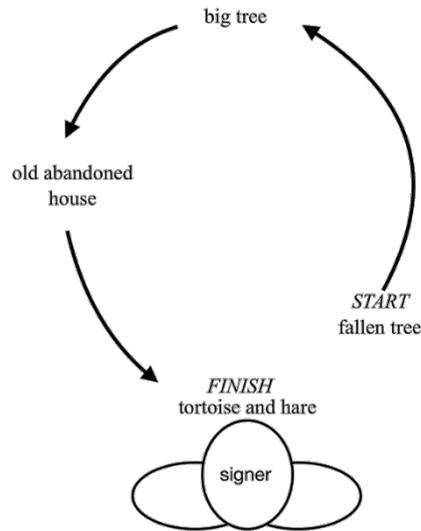
**Figure 6.** Embodiment of the tortoise walking

Italian allows using two different verbs to describe the two actions: “camminare” and “incamminarsi”, respectively. The former is an activity; the latter implies an accomplishment and it is an ingressive verb<sup>4</sup>. It follows that we had to deal with the latter term on an ad hoc basis.

I think that visual modality is what makes sign languages so fascinating. It 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 the generator 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 in which the signer embodies the hare illustrating the route of the race to the tortoise. The route is depicted in Figure 7 on the next page.

---

<sup>4</sup> An ingressive verb indicates that the action is about to be entered.



**Figure 7.** Route of the race in the fable

Segments covered by the arrows in Figure 7 are parts of space indicated by the signer using the index finger with a continuous movement.

I had to gloss these signs 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 sign in Figure 8 were used for a specific section of the route, the character would probably want to specify to his interlocutor that on that section they have to walk, instead of running.



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

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.

We managed to generate correct sentences by glossing these continuous index signs as follows: INIZIARE-GIRO, CONTINUARE-GIRO, FINIRE-GIRO, respectively.

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

In the target translation, verbs are inflected not only for present and past tense, but also for future tense and imperative mood. Therefore, we wanted the generator to be able to produce all of them too.

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 manual signs expressing future tense, but who is watching the fable can easily infer that the mentioned actions are intended to take place in the future.

The generation of this tense did occur in one case, but not in the others.

Unfortunately, we have not been able to find a well-defined linguistically motivated trigger to generate future 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: by imposing the illocutive 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, since these are not illocutive verbs. Thus, they appear in the output text in present and not in future tense.

## 4.4 Summary

In this Chapter, the reader has been provided with an overview of GENLIS.

As explained in Section 4.2, our generator is a machine that receives as input LIS glosses and semantic forms written in Prolog and it is based on particular algorithms. Its core is the Generation algorithm, a grammar algorithm with peripheral rules for exceptions, organised in specific steps. In the above mentioned section we also mapped verbal tense and mood and introduced other two algorithms: the ADA, which is responsible for definiteness assignment, and the Algorithm for narrative direct speech speaking verb type, which allows generation of direct speech.

Section 4.3 was dedicated to the main problems encountered in the generation, which led us to draw some important conclusions. On one hand, they showed that further studies on the organisation 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 glosses. On the other hand, the existence of particular issues in the generation shows that LIS is a real language with specific rules and characteristics and that it does not depend on Italian language. Moreover, from these issues we can also infer that pragmatics and commonsense knowledge play a vital role not only in vocal languages, but in Italian Sign Language too.



# Conclusions

I started this work wondering if it was possible for us to generate an Italian text from Italian Sign Language glosses. We have shown that we have been able to do it, although there still are problems that need to be solved.

After having provided a solid theoretical background on Italian Sign Language and Natural Language Generation, we dove into our experiment, trying to generate Italian from LIS glosses. At this point, we can claim that we achieved our goal. In fact, we have been able to make GENLIS produce an all in all correct Italian text. As we have seen, it does not differ substantially from the target translation in content, but they differ rather in form. In fact, the generated output is an all in all correct text: it is readable and quite detailed compared to the target translation. However, the structure of the output text appears to be rigid and repetitive at some points and some mistakes are also present: agreement errors, mistakes in the use of direct speech and in the creation of anaphoric relations are some of the main problems. Furthermore, tenses are not always selected correctly. In this respect, it is also important to point out that there are no overt past-temporal references in the source video. Nevertheless, we decided to generate using past tense, where possible. This is due to the domain of the text, i.e. the domain of fables, which in Italian are usually told using past tense.

In this work, we also presented the GENLIS generator, which produced the output text that we have analysed and which is however still in an early stage of development. In fact, as shown in Chapter 4 it is organised into some general steps that include a Text To Speech system, which is currently underway and not available yet.

Anyway, generation would not be possible without input information, which in our case derived from LIS manual glosses. As to glosses, I organised them into tables that consist of eight layers (affective, adverbial and syntactic Non-Manual Markers, spatial agreement, Non-Manual Signs, Manual Signs, Action Role Shift and Quotation Role Shift), which contain different linguistic information. From glosses it is evident that one of the peculiarities of sign languages is simultaneity. Simultaneity is an actual problem for generation, since computers operate with sequential strings. Besides, for this reason it was necessary to convert manual glosses into sequential semantic forms written in

Prolog, so that the generator would have been able to process them. However, the main crux here is that simultaneity is difficult to put into writing in an appropriate way. It is not easy to transpose pieces of SL information into different layers of glosses that have specific functions and that have then to be melted together into one string in order to be processed.

Since I noticed that some mistakes in the generated text are caused by the way signs were annotated, I proposed some modifications to glosses. More precisely, I considered the possibility of expanding the Manual Signs tier, in which I annotated manual components. In fact, a sign alone, e.g. a classifier, may convey simultaneous pieces of information, each of which may be expressed by one of the four integral parts of the signs, that are cheremes. Therefore, the idea is to divide the MS layer into four tiers, corresponding to handshape, movement, location and palm orientation of the sign respectively. By doing so, it should be easier to deal with classifiers used in the fable that we glossed. It would be interesting to verify whether this sub-categorisation could be applied to the annotation of other classifiers. However, the use of other annotation systems could also be taken into consideration.

To sum up, this thesis showed the following:

- i. Italian Sign Language is a complex natural language, equal to spoken languages.
- ii. It is possible to generate Italian from Italian Sign Languages glosses, although several issues need to be solved.
- iii. Further research is needed:
  - a. LIS, as a sign language, is a field that only waits to be further explored. For example, further investigations should be conducted on Role Shift and Non Handed Signs. The use of indexical and anaphoric temporal expressions and the role of spatio-temporal coordinates of the narrator when telling a story should be further analysed too.
  - b. The annotation system used for this work needs to be perfected. Then, it is necessary to verify whether it is appropriate for the annotation of other fables and, more broadly, for LIS texts. Other annotation and writing systems should also to be tested.

- c. GENLIS is still in an early stage of development and needs to be optimised. The Text To Speech system should be implemented and optimised too.





**DU 2**

\* = LOC 2\_NPROX

AFF	allezoso		
ADV			
SYN			
AGR	*	*	*
NMS		GUARDARE	
MS	LEPRE: SALTARE I VICINO IX-3*. COMINCIARE: [ ]	CHIEDERE: NOI-2 GARA PUÒ.	
ARS	[ ]	[ ]	[ ]
QRS			

**DU 3**

AFF		perplesso
ADV		
SYN		wh
AGR	LOC7	PROX
NMS	GUARDARE	
MS	TARTARUGA; GARA COSA.	
ARS	[T	]
ORS	[T	]

**DU 4**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	GARA CORRERE ARRIVARE PRIMO CIL.
ARS	
ORS	L. _____

**DU 5**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	TAR TARIQA: VA-BENE, IX-1 GARA ACCETTARE.
ARS	T
ORS	T

**DU 6**

AFF	
ADV	
SYN	WH
AGR	
NMS	
MS	FARE-GIRO COME.
ARS	
ORS	T

**DU 7**

AFF	
ADV	
SYN	topic
AGR	LOC3_NPROX
NMS	LOC3_NPROX
MS	LEPRE: PARTIRE, VEDERE IX-LOC3_PROX ALBERO-CAIDUTO CL, IX-LOC3_PROX PARTIRE IX-LOC3_PROX.
ARS	
QRS	L

**DU 7**

AFF	
ADV	
SYN	
AGR	LOC4_DIST
NMS	LOC4_DIST
MS	POI COMINCIARE CORRERE, L INIZIARE-GIRO IX-LOC4_DIST ALBERO GROSSO IX-LOC4_DIST,
ARS	
QRS	L

**DU 7**

AFF	
ADV	
SYN	
AGR	LOC5_DIST
NMS	LOC5_DIST
MS	POI CONTINUARE-GIRO IX-LOC5_DIST CASA IX-LOC5_DIST VECCHIO ARBANDONATO IX-LOC5_DIST, POI FINIRE-GIRO IX-LOC3_PROX.
ARS	
QRS	L

**DU 8**

AFF	
ADV	
SYN	y/n
AGR	
NMIS	
MS	TAR TARIQA: VA-BENE, GARA ADESSO.
ARS	
QRS	T

**DU 9**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	NO++ , DOMANI MATTINA.
ARS	
QRS	L

**DU 10**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	SIT+, VA BENE.
ARS	
QRS	T

**DU 10**

AFF	
ADV	
SYN	y/h
AGR	LOC3_NPROX
NMS	
MS	APPUNTAMENTO IX-LOC3_PROX.
ARS	
QRS	T

DU 11

AFF	
ADV	
SYN	
AGR	__LOC3_NPROX
NMS	
MS	SI+I, IX-LOC3_PROX.
ARS	
QRS	L

**DU 12**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	BENE, GRAZIE.
ARS	
ORS	T

**DU 13**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	TAR TARIUGA REFERENTE-VA-DX _CL. LEPRE ZAMPETTARE _CL. REFERENTE-VA-SX _CL.
ARS	
ORS	

**DU 14**

\* = LOC6\_NPROX

AFF	_____	entusiasta
ADV	_____	
SYN	_____	
AGR	_____*	
NMS	_____	
MS	ANIMALE BOSCO TUTTI VEDERE. SAPERE DOMANI MATTINA PRESTO ACCORRERE. CI VEDERE+ ASPETTARE ACCORRERE. CL.	
ARS	_____	
ORS	_____	

**DU 15**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	LEFRE PRIMO VENIRE.
ARS	
ORS	

**DU 15**

\* = LOC2\_NPROX

AFF		alCZOS0
ADV		
SYN	FEIOTICA	FEIOTICA
AGR		*
NMS		
MS	SALTARE L. TARTARUGA NON C'È. IX-3T PAURA RINUNCIARE.	
ARS	L	
ORS	L	

**DU 16**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	POI PIANO-PIANO TARTARUGA VENIRE.
ARS	
QRS	

**DU 16**

AFF	
ADV	
SYN	y/n
AGR	
NMS	
MS	CAMMINARE T IX-1 PAURA RINUOCIARE NO 4 IX-1 GARA PARTECPARE IX-1B.
ARS	T
QRS	

**DU 17**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	VA-BENE.
ARS	
QRS	L ]

**DU 18**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	REFERENTE1+REFERENTE2-SI-AFFIANCANO _CL.
ARS	
ORS	

**DU 18**

AFF	
ADV	
SYN	
AGR	LOC2_NPROX
NMS	__SBUFFARE
MS	REFERENTE1 _CL.IX-3I.LEPRE.ZAMPEL _CL.ORECCHIE L.ORECCHIE 1-VS-DIETRO.ZAMPEL _CL.
ARS	
ORS	

**DU 18**

\* = PREPARARSI-ALLA-PARTENZA

AFF	
ADV	
SYN	
AGR	LOC7_PROX
NMS	
MS	TARTARUGA.TRANQUILLLO.GUARDARE-VS-LOC7_PROX.GUARDARE-AVANTI.ZAMPEL _CL.
ARS	
ORS	

**DU 19**

AFF			
ADV			
SYN	wh	y/n	foc
AGR			
NMS			
MS	VENIRE CHI. GIU'FO. VOI-DUE PRONTO. COMINCIARE. 3. 2. 1. VIA.		
ARS			
QRS	G		

**DU 20**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	LEPRE REFERENTE1+REFERENTE2-FERMI_CL REFERENTE1-SUPERAVVELOCIMENTO-REFERENTE2_CL ANDARSENE.
ARS	
ORS	

**DU 20**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	TARTARUGA, UGUALE, CAMMINARE, T.
ARS	T
ORS	

**DU 21**

AFF	
ADV	veloce
SYN	
AGR	
NMS	
MS	LEPRE REFERENTE-CORRE CL, CORRERE L, ORECCHIE L-IN-MOVIMENTO, CORRERE L,
ARS	L
ORS	

**DU 21**

\* = aria di sufficienza

AFF	*
ADV	
SYN	
AGR	
NMS	
MS	GUARDARE-INDIETRO: TARTARUGA IN-FONDO(DIETRO), PUL.
ARS	L
ORS	

DU 22

AFP			
ADV		veloce	
SYN			
AGR			LOC8_NPROX LOC8_NPROX
NMS			
MIS	CORRERE_L. ALBERO GROSSO (ALBERO GROSSO_CL) CORRERE_L. REFERENTE-IN:MOVIMENTO_CL ARRIVARE CASA VECCHIO ABBANDONATO IX:LOC8_NPROX (CASA_CL REFERENTE-IN:MOVIMENTO_CL)		
ARS	L	L	
QRS			

**DU 23**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	SALTARE. L. PU2. TARTARUGA LONTANO. IX-1 COMINCIARE ADDORMENTARSI UN-PO'. PU2.
ARS	L
ORS	L

**DU 23**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	CORRERE. L. (CASA. CL. CASA-DIETRO-ANDARE. CL). DORMIRE-ADDORMENTARSI ORECCHIE. L-ABBASSARSI. DORMIRE.
ARS	L     L
ORS	

**DU 24**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	TARTARUGA TRANQUILLO CAMMINARE T.
ARS	T
ORS	

**DU 24**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	ALBERO GROSSO CAMMINARE T. (ALBERO CL REFERENTE-INMOVIMENTO-VICINO-ALBERO CL), CAMMINARE T. ((ALBERO CL REFERENTE-ALBERO-SUPERARE CL), REFERENTE-ALBERO-SUPERATO CL).
ARS	T
ORS	T

**DU 25**

\* = LOC9\_DIST

AFF				
ADV				
SYN			y/n	
AGR		LOC8_NPROX	*	*
NMS				
MS	COMINCIARE CASA VECCHIO ABBANDONATO IX-LOC8_NPROX CAMMINARE T_VEDERE LEPIRE DORMIRE IX-LOC9_DIST_PU3.			
ARS		T		
ORS			T	

DU 26

AFF			
ADV			
SYN			
AGR		LOC3 PROX	LOC3 PROX
NMS			
MS	CAMMINARE_T (ALBERO_CL REFERENTE-INMOVIMENTO-SUPERA-ALBERO_CL), ARRIVARE QUASI VICINO ALBERO_CADUTO_CL IX-LOC3 PROX, CAMMINARE_T.		
ARS	T		T
QRS			

DU 27

AFF	
ADV	
SYN	wh
AGR	LOC8_PROX
NMS	LOC10_NDIST
MS	IX-LOC8_NPROX LEPRE DORMIRE. SVEGLIARSI. STRACCIARSI. TAR TARUGA DOVE. ANCORA IX-LOC10_NDIST. VEDERE+.
ARS	L
QRS	L

DU 27

AFF	
ADV	
SYN	
AGR	LOC11_NPROX LOC12_DIST LOC12_DIST LOC13_DIST
NMS	
MS	GUARDARE NON-C'E. IN-FONDO GUARDARE NON-C'E. GUARDARE-ALTROVE.
ARS	L L L L
QRS	L L L L

DU 27

AFF	i
ADV	sorpreso
SYN	
AGR	
NMS	
MS	ARRIVARE ALBERO_CADUTO_CL_VICINO CAMMINARE_T.
ARS	
QRS	L

**DU 28**

AFF	
ADV	_____ veloce
SYN	
AGR	
NMS	
MS	LEPRE COMINCIARE CORRERE CORRERE L.
ARS	L_____
ORS	

**DU 29**

AFF		isopreso
ADV		
SYN		
AGR	_LOC13_ DIST	
NMS		
MS	IX-LOC13_ DIST TARTARUGA GUARDARE-INDIETRO.	GIUSCIO-ALZARE_CL_ CAMMINARE-VELOCEMENTE_T_
ARS		[T
ORS		

**DU 30**

AFF	
ADV	
SYN	
AGR	
NMS	
MS	{REFERENTE2_CL,REFERENTE1_RINCORRE_REFERENTE2_CL,REFERENTE1_RAGGIUNGE_REFERENTE2_CL}, PRIMO TARTARUGA, PRIMO {REFERENTE1+REFERENTE2-AL-TRAGUARDO_CL} NASTRO SUPERARE, FINITO.
ARS	[T] [L] [L] [T]
QRS	

**DU 30**

AFF	CON VERGOGNA
ADV	
SYN	
AGR	
NMS	
MS	LEPRE SALTARE_L,ARRIVARE VERGOGNA CAMMINARE_L,ANDARSENE.
ARS	[L] [L]
QRS	

## Appendix B

1

In un bosco viveva una lepre, una lepre altezzosa.

1

Mentre saltellava con aria superba, improvvisamente vide una tartaruga che camminava tranquilla, lentamente.

2

La lepre le si avvicinò saltellando e le chiese con tono sprezzante: "Noi due possiamo fare una gara?"

3

La tartaruga perplessa domandò: "Che gara"?

4

"Correre una gara a chi arriva prima", rispose la lepre.

5

La tartaruga rispose: "Va bene, accetto la gara".

6

Che giro faremo?

7

La lepre rispose: "Per partire, vedi quell'albero caduto? Partiremo da lì"

7

"Poi cominciamo a correre e facciamo il giro del grosso albero in fondo."

7

Poi continuiamo a correre fino alla vecchia casa abbandonata e quindi arriveremo qua.

8

La tartaruga rispose: "Va bene. La gara è adesso?"

9

No, domani mattina", replicò la lepre.

10

"Sì, sì, va bene", disse la tartaruga.

10

"L'appuntamento è qui?", chiese la tartaruga.

11

"Sì sì, qui", disse la lepre.

12

"Perfetto, grazie", disse la tartaruga.

13

La tartaruga se ne andò/va a destra e la lepre zampettando a sinistra.

14

Tutti gli animali del bosco che hanno visto, sapevano che domani mattina presto sarebbero (potuti accorrere) accorsi entusiasti per aspettare e vedere la gara.

15

La lepre venne per prima

15

saltellando, e disse sprezzante: "La tartaruga non c'è? Ha paura e rinuncia?"

16

Poi, lentamente, la tartaruga arrivò:

16

Camminando: "Io rinunciare per paura? No. Io alla gara ci partecipo".

17

"Va bene", rispose la lepre.

18

Le due si affiancarono.

18

La lepre era a sinistra, sbuffava tanto, le orecchie all'indietro, guardando la tartaruga.

18

La tartaruga guardava tranquillamente la lepre e si preparò alla partenza.

La tartaruga guardava tranquilla a sinistra la lepre.

19

Chi arriva ora? Un gufo. "Siete pronte? Cominciamo! 3, 2, 1 ... Via"!

20

La lepre parte come un fulmine e scompare.

20

E anche la tartaruga si incammina.

21

La lepre correva e correva velocemente, e le sue orecchie vibravano.

21

Guardò indietro e disse sprezzante: "La tartaruga è in fondo".

22

Corse fino al grosso albero, lo superò, corse e arrivò alla vecchia casa abbandonata,

23

Saltellando disse: "Beh, la tartaruga è lontana. Mi metto a dormire un po', dai".

23

Di corsa va dietro alla casa, abbassa le orecchie per addormentarsi e dorme.

24

La tartaruga tranquilla camminava.

24

Camminando, si avvicina al grosso albero e lo supera.

25

arriva alla vecchia casa abbandonata e vede la lepre dormire. "Vabbeh", commenta.

26

Continua a camminare e arriva quasi vicino all'albero caduto.

27

La lepre dormiva, si risvegliò e si stiracchiò e disse: "Dov'è la tartaruga? È ancora in fondo? Vediamo ...

27

Lì non c'è, lì in fondo non c'è ...

Guardò da un lato. "Non c'è, e in fondo non c'è". Guardò altrove.

27

Oh, è arrivata vicino all'albero caduto!"

28

La lepre si mise a correre a più non posso.

29

La tartaruga guardò indietro. "Oh", esclamò. Alzò il suo guscio e proseguì più velocemente.

30

La lepre rincorre la tartaruga e la raggiunge, ma la tartaruga taglia il traguardo per prima e vince.

30

La lepre piena di vergogna se ne andò.

## Appendix C

1  
nel bosco viveva una lepre la lepre altezzosa

1  
mentre saltellava vide improvvisamente una tranquilla tartaruga  
che camminava lentamente

2  
la lepre si avvicina saltellando e chiede con un tono  
sprezzante : noi 2 possiamo fare una gara ?

3  
la tartaruga chiede con aria perplessa : quale gara è ?

4  
la lepre replica: " correre la gara a chi arriva prima

5  
la tartaruga dice: " va bene accetto la gara

6  
quale giro faremo ?

7  
la lepre risponde: " per partire guarda un albero caduto  
partiremo qua

7  
poi cominciamo a correre e fa il giro

7  
poi continua a correre fino a la vecchia casa abbandonata  
quindi finisce il giro qua

8  
la tartaruga risponde: " va bene la gara è adesso ?

9  
la lepre risponde: " no domani mattina

10  
la tartaruga risponde: " sì sì va bene

10  
la tartaruga dice: " : l appuntamento è qui ? ?

11  
la lepre risponde: " sì sì qui

12  
la tartaruga replica: " va bene grazie

13  
la tartaruga se ne andò a destra e la lepre se ne andò  
zampettando a sinistra

14  
tutti gli animali che avevano visto sanno che sarebbero accorsi domani mattina presto per aspettare con spirito entusiasta per vedere la gara con spirito entusiasta

15  
la lepre venne per prima

15  
saltellava

e dice con aria sprezzante la tartaruga non è qui  
ha paura e rinuncia ?

16  
poi la tartaruga venne piano piano

16  
chiede : rinuncia per paura ?

16  
risponde: " no dice a la gara ci partecipava

17  
la lepre replica: " va bene

18  
i due si affiancarono

18  
la lepre stava zampettando sbuffava tanto a guardare la tartaruga

18  
la tartaruga guardava a sinistra tranquilla la lepre zampettare

19  
ora arriva un gufo e dice : voi due siete pronti ?

19  
cominciamo dice: " 3 2 1 ... via

20  
la lepre superò la tartaruga velocemente e se ne andò

20  
e anche la tartaruga si incamminò

21  
la lepre correva

e correva velocemente le orecchie vibravano

21  
la lepre guardò indietro e disse sprezzante la tartaruga è dietro in fondo

22

poi correva fino a l albero grosso

sorpassò l albero grosso

22

correva e arrivò a la vecchia casa abbandonata

23

saltellava

e dice la tartaruga è lontana io comincio a dormire un pò  
Dai

23

andò di corsa dietro la casa e abbassò le orecchie per dormire

23

e si addormentò

24

la tartaruga camminava tranquilla

24

si avvicinava camminando a l albero grosso e superò l albero

25

arrivò accanto a la vecchia casa abbandonata

25

vede la lepre dormire e dice Vabbeh

26

la tartaruga camminava e è arrivata quasi vicino a l albero caduto

27

la lepre dormiva

si svegliò e si stiracchiò

27

dice: " : dove è la tartaruga ? : è ancora là ? ? vedremo

27

guardava di lato e non è in fondo

27

guardò altrove

27

dice: " Oh la tartaruga è arrivata quasi vicino a l albero caduto

28

la lepre corre a più non posso

29

la tartaruga guardava indietro e esclama Oh

29

alzò il guscio energicamente e camminò velocemente

30

la lepre rincorreva la tartaruga e la raggiunse

30

ma la tartaruga tagliò il traguardo per prima e vinse

30

la lepre arrivò saltellando e se ne andò con vergogna



## References

- Aarons, D. R., Morgan, G. (2003), "Classifier Predicates and the Creation of Multiple Perspectives in South African Sign Language". In: *Sign Language Studies* 3, (125-56).
- Ahlgren, I. (1990), "Deictic pronouns in Swedish and Swedish Sign Language". In S. D. Fischer, P. Siple (eds.), *Theoretical Issues in Sign Language Research Vol. 1*, (pp. 167-174). Chicago: University of Chicago Press.
- Aikhenvald, A. Y. (2000), *Classifiers: A Typology of Noun Categorization Devices*. Oxford University Press, Oxford.
- Ajello, R. (1997), "Lingue vocali, lingue dei segni e 'l'illusion mimétique'". In: R. Ambrosini, P. Bologna, F. Motta, C. Orlandi (eds.), *Schribthair a anim n-ogaim*, (pp. 17-30), Pisa, Pacini Editore.
- Allan, K. (1977), "Classifiers". In: *Language*, 53 (pp. 285-311).
- Appelt D. (1985), *Planning English Sentences*, Cambridge University Press, Cambridge, UK.
- Bahan, B. (1996), *Nonmanual Realization of Agreement in American Sign Language*. Doctoral dissertation, Boston University, Boston Mass.
- Bahan, B., Pettito, L. (1980), *Aspects of Rules for Character Establishment and Reference in ASL Storytelling*. Unpublished manuscript. Salk Institute for Biological, Studies, La Jolla, CA.
- Baker, C., Padden, C. (1978), "Focusing on the non-manual components of ASL". In P. Siple (ed.), *Understanding language through sign language research*. (pp 27-57). New York, Academic Press.
- Baker-Shenk, C. (1983), *A microanalysis of the non-manual components of questions in American Sign Language*. Unpublished doctoral dissertation, Berkeley, University of California.
- Barberis D., Garazzino N., Prinetto P., Tiotto G., Savino A., Shoaib U., Ahmad N. (2011), "Language Resources for Computer Assisted Translation from Italian to Italian Sign Language of Deaf People". In: *Proceedings of Accessibility Reaching Everywhere AEGIS Workshop and International Conference*, Brussels, Belgium (November 2011).

- Battaglino C., Geraci C., Lombardo V. Mazzei A. (2015). “Prototyping and Preliminary Evaluation of Sign Language Translation System in the Railway Domain”. In: M. Antona, C. Stephanidis (eds.), *Universal Access in Human-Computer Interaction. Access to Interaction*, (pp. 339–350), Springer.
- Battison R. (1978), *Lexical borrowing in American Sign Language*. Linstok Press.
- Bauer B., Nießen S., Hienz H. (1999) “Towards an Automatic Sign Language Translation System”. In: *Proceedings of International Workshop on Physicality and Tangibility in Interaction: Towards New Paradigms for Interaction Beyond the Desktop*. Siena, Italy (1999).
- Benedicto E., Brentari D. (2004), “Where Did All the Arguments Go? Argument-Changing Propertiers of Classifiers in ASL”. In: *Natural Language and Linguistics Theory*, 22(4), (pp. 743-810).
- Bertinetto P. M. (1986), *Tempo, Aspetto e Azione nel verbo italiano. Il sistema dell’indicativo*, Firenze, Accademia della Crusca.
- Bertinetto P. M. (1991), “Il verbo”. In: L. Renzi, G. Salvi, A. Cardinaletti (eds.), *La grande grammatica italiana di consultazione*, Bologna, Il Mulino.
- Bertone C. (2011), *Fondamenti di grammatica della Lingua dei Segni Italiana*, Milano, Franco Angeli.
- Bertone C. (2009), “La grammatica dello spazio nella LIS”, in C. Bertone, A. Cardinaletti (eds.), *Alcuni capitoli della grammatica della LIS. Atti dell’Incontro di Studio “La grammatica della Lingua dei Segni Italiana” Venezia, 16-17 maggio 2007*, Venezia, Cafoscarina.
- Bertone C. (2011), *Fondamenti di Grammatica della Lingua dei Segni Italiana*, Milano, FrancoAngeli.
- Binstead K., Ritchie G. (1997), “Computational rules for pruning riddles”. In: *Humor 10*, (pp. 25-76).
- Boyes Braem, P. (1981), *Features of the Handshape in American Sign Language*, Unpublished doctoral dissertation, University of California, Berkeley.
- Boyes Braem, P., Sutton-Spence, R. L. (2001), *The Hands Are The Head of The Mouth. The Mouth as Articulator in Sign Languages*. Hamburg, Signum Press.
- Bowden K. K., Lin G. I., Reed L. I., Fox Tree J. E., Walker M. A. (2017), “M2D: Monolog to Dialog Generation for Conversational Story Telling”. In: F. Nack, A.

- S. Gordon (eds.), *Interactive Storytelling*, volume 10045 of *Lecture Notes in Computer Science*, (pp. 12–24), Los Angeles, CA, USA, Springer International Publishing.
- Branchini, C. (2007), “On relativization and clefting in Italian Sign Language”. In *Sign Language & Linguistics*, vol. 10.2, (pp. 201-212), Amsterdam, John Benjamins.
- Branchini, C., Donati, C. (2009) “Relatively different: Italian Sign Language relative clauses in a typological perspective”. In: Liptak, A. (ed.), *Correlatives Cross-Linguistically*, 157-191. Amsterdam: Benjamins.
- Branchini, C., Geraci, C. (2011), “L'ordine dei costituenti in LIS: risultati preliminari”. In: A. Cardinaletti, C. Cecchetto, C. Donati, *Grammatica, lessico e dimensioni di variazione nella LIS* (pp. 113-126). Milano, Franco Angeli.
- Branchini, C. (2014), *On relativization and clefting. An analysis of Italian Sign Language*, Berlin, Mouton De Gruyter.
- Brennan, M. (1994), “Word Order: Introducing the Issues”. In: M. Brennan, H. Graham (eds.), *Word Order Issues in Sign Language. Working Papers*, (pp. 9-46), Durham, International Sign Linguistics Association.
- British Deaf Association (1975), *Gestuno: International Sign Language of the Deaf*. Carlisle, British Deaf Association.
- Brown J. C., Frishkoff G. A., Eskenazi M. (2005). “Automatic question generation for vocabulary assessment”. In: *Proc. EMNLP'05*, (pp. 819-826).
- Büring, Daniel (2005), *Binding Theory*. Cambridge, Cambridge University Press.
- Caselli, M. C., Maragna, S., Volterra, V. (2006), *Linguaggio e sordità. gesti segni e parole nello sviluppo e nell'educazione*. Bologna, Il Mulino.
- Cawsey A., Jones R., Pearson J., (2000), “The evaluation of a personalised health information system for patients with cancer”. In: *User Modelling and User-Adapted Interaction 10*, (pp. 47-72).
- Clark, H. H., Gerrig, R. J. (1990), “Quotations as Demonstrations”. In: *Language*, 66(4), (pp. 764-805).
- Cecchetto, C., Geraci, C., Zucchi, S. (2009), “Another way to mark syntactic dependencies. The case for right peripheral specifiers in sign languages”. In: *Language* 85 (2), (pp. 278-320).

- Chen D. L., Mooney R. J. (2008), “Learning to sportscast: a test of grounded language acquisition”. In *Proc. ICML '08*, (pp. 128–135).
- Cheng H., Mellish C. (2000), “Capturing the interaction between aggregation and text planning in two generation systems”. In *Proc. INLG '00, Vol. 14*, (pp. 186–193).
- Chiari I. (2007), *Introduzione alla Linguistica Computazionale*. Roma, Laterza.
- Coch, J. (1998). “Interactive generation and knowledge administration in MultiMeteo”. In *Proc. IWNLG '98*, (pp. 300–303).
- Corazza, S. (1990), “The Morphology of Classifier Handshapes in Italian Sign Language (LIS)”. In: C. Lucas (ed.), *Sign Language Research: Theoretical Issues*, (pp. 71-82), Gallaudet University Press, Washinton, D.C.
- Corina, D. P., Bellugi, U., Reilly, J. (1999) “Neuropsychological studies of linguistic and affective facial expressions in Deaf Signers”. In: *Language and Speech*, 42(2-3), (pp 307-331).
- Cormier, K., Fenlon, J., Rentelis, R., Schembri, A. (2011), “Lexical frequency in British Sign Language conversation: a corpus-based approach”. In: P.K. Austin, O. Bond, L. Marten, D. Nathan (eds.), *Proceedings of the Conference on Language Documentation and Linguistic Theory 3*, (pp. 81-90), London, School of Oriental and African Studies.
- Cormier, K., Schembri, A., Woll B. (2013), “Pronouns and pointing in sign languages”. In *Lingua 137*, (pp 230-247). Elsevier.
- Craig C. (1986), *Noun classes and categorization*. Hohn Benjamins, Amsterdam.
- Craig C. (1992), “Classifiers in a functional perspective”. In: M. Fortescue, P. Harder, L. Kristoffersen (eds.), *Layered structure and reference in a functional perspective*, (pp. 277-301). John Benjamins, Amsterdam.
- Crasborn, O. A. (2015), “Transcription and Notation Methods”. In: *Research Methods in Sign Language Studies: A Practical Guide*, (pp.74-88), London, Wiley Blackwell.
- Cunningham P., Veale T. (1991), “Organizational Issues Arising from the Integration of the Concept Network and Lexicon in a Text Understanding System”. In: *12th International Joint Conference on Artificial Intelligence, Sydney, Australia*, (pp. 1009–1015).

- Cuxac, C., 2000. "Iconicity of Sign Language". In: M. Taylor, F. Néel, D.G. Bouwhuis (eds.), *The structure of multimodal dialogue*. Benjamins, Amsterdam, (pp. 321-334).
- Dale R., Anisimoff I., Narroway G. (2012). "Hoo 2012: A report on the preposition and determiner error correction shared task". In: *Proc. 7th Workshop on Building Educational Applications Using NLP*, (pp. 54-62).
- Denny, J. P. (1976), "What's the Use of a Classifier?". In: *Papers from the Annual Regional Meeting of the Chicago Linguistic Society, 12*, (pp. 122-132).
- Di Renzo, A., Lamano, L., Luciola, T., Pennacchi, B., Gianfreda, G., Petitta, G., Bianchini C. S., Rossini, P., Pizzuto E. (2011), *Scrivere la LIS con il Sign Writing. Manuale introduttivo*, Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Roma.
- Dively, V. L. (2001), "Signs without Hands: Nonhanded Signs in American Sign Language". In: V. Dively, M. Metzger, S. Taub and A. M. Baer (eds.), *Signed Languages – Discoveries from International Research*. (pp. 62-73), Washington DC, Gallaudet University Press.
- Dixon, R. M. W. (1968), "Noun classes". In: *Lingua, 21*, (pp. 104-125)
- Dixon, R. M. W. (1982), *Where Have All the Adjectives Gone? And Other Essays in Semantics and Syntax*. Mouton Publishers, Berlin.
- Donati, C., Barbera, G., Branchini, C., Cecchetto, C., Geraci, C., Quer, J. (2017), "Searching for imperatives in European sign languages". In: *Imperatives and Directive Strategies*, vol. 184, (pp. 111-155). Amsterdam/Philadelphia, John Benjamins Publishing Company.
- Dreuw P., Forster J., Gweth Y., Stein D., Ney H., Martinez G., Llahi J.V., Crasborn O., Ormel E., Du W., Hoyoux T., Piater J., Moya J.M., Wheatley M. (2010), "Signspeak – Understanding, Recognition, and Translation of Sign Languages". In: *Proc. of 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, (pp. 22-23).
- Dryer, M. S. (1989), "Discourse-Governed Work Order and Work Order Typology". In: *Belgian Journal of Linguistics 4* (pp. 69-90).

- Dryer, M. S. (2007), "Word Order". In: T. Shopen (ed.), *Language Typology and Syntactic Description. Vol. 1: Clause Structure* (2<sup>nd</sup> Edition) (pp. 61-131), Cambridge, Cambridge University Press.
- Earis, H., Cormier, K. (2013), "Point of view in British Sign Language and spoken English narrative discourse: The example of "The Tortoise and the Hare".". In: *Language and Cognition*,5(4), (pp. 313-343).
- Ebling S. (2016), *Automatic Translation from German to Synthesized Swiss German Sign Language*, PhD Dissertation, University of Zurich.
- Efthimiou E., Fotinea S.-E., Hanke T., Glauert J., Bowden R., Braffort A., Collet C., Maragos P., Goudenove F. (2010), "DICTA-SIGN: Sign Language Recognition, Generation and Modelling with application in Deaf Communication". In: *Proceedings of CSLT 2010 (LREC 2010)*, (pp. 80-83).
- Elliott R., Glauert J.R.W., Kennaway J.R., Marshall I. (2000), "The Development of Language Processing Support for the ViSiCAST project". Presented at *4th International ACM SIGCAPH Conference on Assistive Technologies (ASSETS 2000)*.Washington.
- Emmorey, K. (1999), "Do signers gesture?" In Lynn S. Messing & Ruth Campbell (eds.), *Gesture, speech, and sign*, (pp 133-159). Oxford, Oxford University Press.
- Emmorey, K., Reilly, J. S. (1998) "The development of quotation and reported action: Conveying perspective in ASL". In: E. Clark (ed.), *The Proceedings of the Twenty-ninth Annual Child Language Research Forum*, (pp. 81-90. CSLI Publications, Stanford, CA.
- Engberg-Pedersen, E. (1993), *Space in Danish Sign Language: The Semantics and Morphosyntax of the Use of Space in a Visual Language*. Hamburg, Signum.
- Engberg-Pedersen, E. (1995), "Point of view expressed through shifters" In: K. Emmorey, J. S. Reilly (eds.), *Language, Gesture, and Space*, (pp. 133-154), Lawrence Erlbaum, Hillsdale, NJ.
- Espinosa D., White M., Mehay D. (2008). "Hypertagging: Supertagging for surface realization with CCG". In: *Proc. ACL-HLT'08*, (pp. 183–191), Columbus, Ohio.
- Ferrari G. (1991), *Introduzione al Natural Language Processing*. Bologna, Calderini.
- Ferres L., Parush A., Roberts S., Lindgaard G. (2006), "Helping people with visual impairments gain access to graphical information through natural language: the

- iGraph system”. In: *Proceedings of the 10th International Conference on Computers Helping People with Special Needs*, (pp. 1122-30).
- Franchi, M. L. (1987), “Componenti non manuali”. In: V. Volterra (ed.), *La lingua dei segni italiana*, (pp 159-177). Bologna: Il Mulino. (Nuova Edizione 2004).
- Friedman, L. A. (1975), “Space, time, and person reference in American Sign Language”. In: *Language* 51(4), (pp. 940–961).
- Frishberg, N. (1975), “Arbitrariness and Iconicity: Historical Change in American Sign Language”. In: *Language*, 51, (pp 696-719).
- Frishberg, N., Hoiting, N., Slobin, D. I. (2012), “Transcription”. In: Pfau, R., Steinbach, M., Woll, B. (eds.), *Sign Language. An International Handbook*, (pp. 1045-1075), Berlin, De Gruyter.
- Gal A., Lapalme G., Saint-Dizier P., Somers H. (1991), *Prolog for Natural Language Processing*. Chinchester, John Wiley & Sons Ltd.
- Garcia, B., Sallandre, M. A. (2013), “Transcription systems for sign languages: a sketch of the different graphical representations of sign language and their characteristics”. In: C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, S. Tessendorf, *Handbook “Body-Language-Communication”*, (pp. 1125-1338), Mouton de Gruyter.
- Gatt A., Belz A., Kow E. (2009), “The TUNA-REG challenge 2009: overview and evaluation results”. In: *Proceedings of the 12th European Workshop on Natural Language Generation (ENLG-2009)*, (pp. 1174-1182).
- Gatt A., Krahmer E. (2017), “Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation”. In: *Journal of Artificial Intelligence Research* 61 (10.1613/jair.5714).
- Geraci, C. (2002), *L'ordine delle parole nella LIS*, Dissertation, Università Statale di Milano.
- Geraci, C. (2006), “Negation in LIS (Italian Sign Language)” In: L. Bateman, C. Ussery (eds.), *Proceedings of NELS 35*, (pp. 217-229). GLSA, Amherst, MA.
- Geraci, C., Cecchetto, C., Zucchi, S. (2008) “Sentential complementation in Italian Sign Language”. In: M. Grosvald, D. Soares (eds.), *Western Conference On Linguistics (WECOL)*, (pp. 46-58), Davis, Department of Linguistics, University of California at Davis.

- Goldberg E., Driedger N., Kittredge R. I. (1994). "Using Natural Language Processing to Produce Weather Forecasts". In: *IEEE Expert*, 2, (pp. 45–53).
- Greenberg, J. (1966), *Universals of Language*. Cambridge, Mass., MIT Press.
- Grieve-Smith A. B. (1999), "English to American Sign Language Machine Translation of Weather Reports". In: *Proceedings of the Second High Desert Student Conference in Linguistics (HDSL2)*, (pp. 13–30), Albuquerque, NM.
- Grimes, J. E. (1975), *The Thread of Discourse*. Seria Minor, Mouton.
- Grosso, B. (1992-1993), *Iconicità ed arbitrarietà nella lingua dei segni italiana. Uno studio sperimentale*. Tesi del Corso di Laurea in Psicologia, Università di Padova.
- Grosso, B. (1997). Gli udenti capiscono i segni? In M.C. Caselli e S. Corazza (eds.), *LIS. Studi, esperienze e ricerche sulla lingua dei Segni in Italia. Atti del 1° Convegno Nazionale sulla Lingua dei Segni. Trieste 13-15 ottobre 1995*. (pp. 79-86) Tirrenia (Pisa): Edizioni del Cerro.
- Hawkins, J. (1983), *Word Order Universals*, New York, Academic Press.
- Herrmann, A. (2013), *Modal and Focus Particles in Sign Languages: A Cross-Linguistic Study*. Berlin, Boston, De Gruyter Mouton.
- Herrmann, A., Steinbach, M. (2012), "Quotation in Sign Languages. A Visible Context Shift". In: I. Buchstaller, I. Van Alphen (eds.), *Quotatives. Cross-linguistic and Cross-disciplinary Perspectives*, (pp. 203-228). Amsterdam: Benjamins.
- Herrmann, A., Pendzich, N.-K. (2018), "Between narrator and protagonist in fables of German Sign Language". In: A. Hübl, M. Steinbach (eds.), *Linguistic foundation of narration in spoken and sign languages* (pp. 275-308), Amsterdam / Philadelphia, John Benjamins Publishing Company.
- Hickok, G., Bellugi, U., Klima, E. S. (1996) "The neurobiology of sign language and its implication for the neural basis of language". In: *Nature*, 381, (pp 699-702).
- Hoiting, N., Slobin, D. I. (2002), "Transcription As A Tool For Understanding: The Berkeley Transcription System For Sign Language Research (BTS)". In: G. Morgan, B. Woll (eds.) (2002), *Directions in sign language acquisition* (pp. 55-75). Amsterdam/Philadelphia: John Benjamins.
- Hovy E. H. (1988), *Generating Natural Language Under Pragmatic Constraints*, Hillsdale, NJ, Lawrence Erlbaum.

- Hovy E. H. (1993), “Automated Discourse Planning and Generation”. In: *Proc. Annual Meeting of the Society for Text and Discourse*.
- Huenerfauth M. (2006), *Generating American Sign Language Classifier Predicates for English-to-ASL Machine Translation*. PhD thesis, University of Pennsylvania, Philadelphia, PA.
- Hüske-Kraus D. (2003), “Suregen 2: a shell system for the generation of clinical documents”. In: *Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL-2003)* (Research Notes and Demos), (pp. 215-218).
- Kazemzadeh S., Ordonez V., Matten M., Berg, T. (2014), “ReferItGame: Referring to Objects in Photographs of Natural Scenes”. In: *Proc. EMNLP’14*, (pp. 787-798).
- Klima, E., Bellugi U. (1979), *The Signs of Language*. Cambridge, Harvard University Press.
- Jacobowitz, E. L., Stokoe W. (1988), Signs of tense in ASL Verbs”, (331–340). In: *Sign Language Studies 60*.
- Kukich K. (1983), “Design and implementation of a knowledge-based report generator”. In: *Proceedings of 21st Annual Meeting of the Association for Computational Linguistics (ACL-1983)*, (pp. 145–50).
- Kulkarni G., Premraj V., Ordonez V., Dhar S., Li S., Choi Y., Berg A. C., Berg T. (2013). “Baby talk: Understanding and generating simple image descriptions”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(12)*, (pp. 2891-2903).
- Langkilde-Geary I. (2000). “Forest-based statistical sentence generation”. In: *Proc. ANLP-NAACL’00*, (pp. 170-177).
- Langkilde-Geary I., Knight K. (2002). “HALogen Statistical Sentence Generator”. In: *Proc. ACL’02 (Demos)*, (pp. 102-103).
- Laudanna, A. (1987), “Ordine dei segni nella frase”. In: *La Lingua Italiana dei Segni. La comunicazione visivo-gestuale dei sordi*. Bologna: Il Mulino.
- Laudanna, A., Volterra V. (1991), “Order of words, signs and gestures: A first comparison”. In: *Applied Psycholinguistics 12*, (pp. 135-150).
- Lee, R. G., Neidle, C. MacLaughlin, D., Bahan, B. Kegl, J. (1997). “Role shift in ASL: A syntactic look at direct speech”. In: C. Neidle, D. MacLaughlin, R. G. Lee

- (eds.), *Syntactic Structure and Discourse Function: An Examination of Two Constructions in American Sign Language*. American Sign Language Linguistic Research Project Report No. 4. Boston MA, Boston University.
- Levelt W. (1989), *Speaking: From Intention to Articulation*, MIT Press, Cambridge, MA.
- Levelt W., Roelofs A., Meyer A. S. (1999), “A theory of lexical access in speech production”. In: *The Behavioral and brain sciences*, 22(1), (pp. 1-38; discussion 38-75).
- Liddell, S.K. (1978), “Non-manual signals and relative clauses in American Sign Language.” In P. Siple (Ed.), *Understanding language through sign language research*. (pp 59-90). New York, Academic Press.
- Liddell, S.K. (1980), *American Sign Language syntax*. The Hague, Mouton.
- Liddell, S. K. (1995), “Real, Surrogate, and Token Space: grammatical Consequences in ASL”. In K. Emmorey, J. S. Reilly (eds.), *Language, Gesture, and Space*, 19-41. Hillsdale, N.J., Lawrence Erlbaum.
- Liddell, S. K. (2002), Indicating verbs and Pronouns, Pointing away from Agreement. In K. Emmorey, H. Lane (eds.), *The Sign of Language Revisited*, (pp 303-320). Hillsdale, NJ, Erlbaum.
- Liddell, S. K., Johnson, R. (1987), *An Analysis of Spatial-Locative Predicates in American Sign Language*. Paper presented at the Fourth International Symposium on Sign Language Research, Lappeeranta, Finland.
- Lillo-Martin, D. (1995), “The Point of View Predicate in American Sign Language”. In: K. Emmorey, J. Reilly (eds.), *Language, Gesture, and Space*, Hillsdale, NJ, Lawrence Erlbaum, (pp. 155-170).
- Lillo-Martin, D. (2012), “Utterance reports and constructed action”. In: R. Pfau, M. Steinbach, B. Woll (eds.), *Sign Language – An International Handbook*, (pp 365-387). Amsterdam, Mouton de Gruyter.
- Lombardo V., Battaglino C., Damiano R. (2011), “An avatar-based interface for the Italian Sign Language”. *International Conference on Complex, Intelligent, and Software Intensive Systems*, (pp. 589-594).

- López-Ludeña V., Barra-Chicote R., Lutfi S., Montero J.M., San-Segundo R. (2013), “LSESpeak: A spoken language generator for Deaf people”. In: *Expert Systems with Applications 40*, (pp. 1283–1295).
- Mac Laughlin, D. (1997), *The Structure of Determiner Phrase: Evidence From American Sign Language*. PhD Dissertation, Boston University.
- Mao J., Huang J., Toshev A., Camburu O., Yuille A., Murphy K. (2016), “Generation and Comprehension of Unambiguous Object Descriptions”. In: *Proc. CVPR’16*.
- Mann W. C., Thompson S. A. (1988). “Rhetorical structure theory: Toward a functional theory of text organization”. In: *Text*, 8(3), (pp. 243–281).
- Marshall I. and Sáfár E. (2003). “A Prototype Text to British Sign Language (BSL) Translation System”. In: *Proceedings of the 41st Annual Meeting of the Association of Computational Linguistics (ACL-03) Conference*, (pp. 113-116), Sapporo, Japan.
- Matthews, P. H. (1974), *Morphology. An introduction to the theory of word-structure*. Cambridge, Cambridge University Press.
- Mazzoni, L. (2008), “Classificatore del corpo e impersonamento in LIS”. In: C. Bagnara, S. Corazza, S. Fontana, A. Zuccalà (eds.), *I segni parlano. Prospettive di ricerca sulla Lingua dei Segni Italiana*, Roma, Francoangeli, pp. (64-75).
- Mazzoni, L. (2009), “Evidenzialità e Impersonamento in LIS”. In: C: Bertone, A. Cardinaletti (ed.), *La grammatica della Lingua dei Segni Italiana. Atti dell’incontro di studio. Venezia 16-17 maggio 2007*, Venezia, Libreria Editrice Cafoscarina.
- Mazzoni, L. (2012), *Classificatori e impersonamento nella Lingua dei Segni Italiana*, Pisa, Pisa University Press.
- McCoy K., Pennington C., Suri L. (1996), “Considering the effects of second language learning on generation”. In: *Proceedings of the Eighth International Workshop on Natural-Language Generation (INLG-1996)*, (pp. 71-80).
- McCullough, S., Emmorey, K., Sereno, M. (2005) “Neural organization for recognition of grammatical and emotional facial expression in deaf ASL signers and hearing nonsigners”. In: *Cognitive Brain Research 22*, (pp 193-203). Elsevier.

- McDonald D. D. (1987), "Foreword". In: Zock M., Sabah G. (eds.) (1988), *Advances in Natural Language Generation. An Interdisciplinary Perspective. Volume 1*. London, Pinter Publishers (pp. ix-xi).
- Macdonald I., Siddharthan, A. (2016). "Summarising news stories for children". In: *Proc. INLG'16*, (pp. 1-10), Edinburgh, UK.
- McKeown, K. R. (1985). *Text Generation*, Cambridge, UK, Cambridge University Press.
- McKeown K., Swartout W. R. (1988), "State of the art". In: Zock M., Sabah G. (eds.), *Advances in Natural Language Generation. An Interdisciplinary Perspective. Volume 1*. London, Pinter Publishers (pp. 1-51).
- Meier, R. P., 1990. "Person Deixis in American Sign Language". In S. D. Fischer, P. Siple (Eds.), *Theoretical Issue in Sign Language Research*. Vol. 1, (pp 175-190). Chicago, University of Chicago Press.
- Mellish C., Scott D., Cahill L., Paiva D. S., Evans R., Reape M. (2006), "A Reference Architecture for Natural Language Generation Systems". In: *Natural Language Engineering*, 12(01), (pp. 1–34).
- Metzger, M. (1995), "Constructed dialogue and constructed action in American Sign Language". In: C. Lucas (eds.), *Sociolinguistics in Deaf Communities*, (pp. 255-271), Washington D.C., Gallaudet University Press.
- Michaud L., McCoy K.F., Pennington C.A. (2000), "An Intelligent Tutoring System for Deaf Learners of Written English". In: *Proceedings of ASSETS 2000, November 13-15*.
- Miller, C. (2006), Sign language: Transcription, notation, and writing. In K. Brown (ed.), *Encyclopedia of Language & Linguistics*, (pp. 353-354). Oxford: Elsevier.
- Mitchell M., van Deemter K., Reiter E. (2013), "Generating Expressions that Refer to Visible Objects". In: *Proc. NAACL'13*, (pp. 1174–1184).
- Molina, M., Stent, A., & Parodi, E. (2011). "Generating Automated News to Explain the Meaning of Sensor Data". In: Gama J., Bradley E., Hollmén J. (eds.), *Proc. IDA 2011*, ( pp. 282-293). Springer, Berlin and Heidelberg.
- Morrissey S. (2008), "Assistive translation technology for deaf people: translating into and animating Irish sign language". In: *ICCHP 2008 - 12th International*

*Conference on Computers Helping People with Special Needs, Young Researchers' Consortium*, (pp. 8-14, Linz), Austria.

Morrissey S. (2011), "Assessing three representation methods for sign language machine translation and evaluation". In: *Proceedings of the European Association for Machine Translation (EAMT)*, (pp. 137-144), Leuven, Belgium.

Morrissey S., Way A. (2013), "Manual labour: tackling machine translation for sign languages". In: *Machine Translation, Vol. 27, No. 1*, (pp. 25-64).

Neidle, C., Kegl, J., Dawn MacLaughlin, Bahan, B., Lee, R. G. (2000). *The syntax of American Sign Language. Functional categories and hierarchical structure*. Cambridge: MIT Press.

Nespor, M., Sandler, W. (1999) "Prosody in Israeli Sign Language". In: W. Sandler (ed.), *Language and Speech* 42(2-3), (pp 143–176).

Padden, C. (1986), "Verbs and role shifting in ASL". In: C. Padden (ed.), *Proceedings of the Fourth National Symposium on Sign Language Research and Training*, Silver Spring MD, National Association of Deaf, (pp. 44- 57).

Padden, C., (1990), "The relation between Space and Grammar in ASL Verb Morphology". In C. Lucas (Ed.), *Sign Language Research, Theoretical Issues*. (pp 118-132). Washington, D.C., Gallaudet University Press.

Paris C., Vander Linden K., Fischer M., (1995), "A support tool for writing multilingual instructions". In: *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-1995)*, (pp. 1398-1404).

Pfau, R., Quer, J. (2010), "Nonmanuals: their grammatical and prosodic roles". In D. Brentari (ed.), *Sign Languages*, (pp. 381-402). Cambridge, Cambridge University Press.

Pietrandrea, P. (2000), "Complessità dell'interazione di iconicità e arbitrarietà nel lessico della LIS." In C. Bagnara, P. Chiappini, M.P. Conte e M. Ott (eds.), *Viaggio nella città invisibile. Atti del 2° Convegno nazionale sulla Lingua Italiana dei Segni. Genova, 25-27 settembre 1998*. (pp. 38-49). Pisa, Edizioni del Cerro.

Pietrandrea, P. (2002), "Iconicity and Arbitrariness in Italian Sign Language". In: *Sign Language Studies*, vol. 2, n° 3, (pp. 296-321).

- Pizzuto, E. (1986), "The verb system of Italian Sign Language". In B. T. Tervoort (ed.), *Signs of life*, (pp 17-31). Amsterdam, University of Amsterdam.
- Pizzuto, E. (1987), "Aspetti morfosintattici", in Volterra V. (ed.), *La lingua italiana dei segni. La comunicazione visivo gestuale dei sordi*. (pp 179-209). Bologna, Il Mulino.
- Pizzuto, E., Giuranna, E., Gambino, G. (1990) "Manual and nonmanual morphology in Italian Sign Language: grammatical constraints and discourse processes". In: C. Lucas (ed.), *Sign Language Research. Theoretical Issues*, (pp 83-102). Gallaudet University Press.
- Pizzuto, E., Cameracanna, E., Corazza, S., Volterra, V. (1995), "Terms for spatio-temporal relations in Italian Sign Language". In: Raffaele Simone (ed.), *Iconicity in language* (pp. 237-256), Amsterdam, John Benjamins.
- Pizzuto, E., Corazza, S. (1996), "Noun morphology in Italian Sign Language (LIS)". In *Lingua 98*, (pp 169-196). Elsevier.
- Pizzuto E., Rossini P. e Russo T. (2006). "Representing signed languages in written form: questions that need to be posed". In: C. Vettori (eds.), *Proceedings of the 2d Workshop on the Representation and Processing of Sign Languages "Lexicografic matters and didactic scenarios" – LREC 2006 - 5th International Conference on Language resources and evaluation, Genova (Italia) 28/05/2006*, (pp. 1-6), Pisa (Italia), ILC-CNR.
- Pizzuto, E., Chiari, I., Rossini P., (2010) "Representing sign language: Theoretical, methodological and practical issues". In: M. Pettorino, A. Giannini, I. Chiari, F. Dovetto (eds.) (pp. 205-240). *Spoken Communication*, Cambridge Scholars Publishing.
- Plachouras V., Smiley C., Bretz H., Taylor O., Leidner J. L., Song D., Schilder F. (2016). "Interacting with financial data using natural language". In: *Proc. SIGIR'16*, (pp. 1121-1124).
- Portet F., Reiter E., Gatt A., Hunter J. R., Sripada S., Freer Y., Sykes C. (2009), "Automatic generation of textual summaries from neonatal intensive care data". In: *Artificial Intelligence, 173(7-8)*, (pp. 789-816).

- Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Henning, J. (1989), *Hamburg Notation System for Sign Languages. An Introductory Guide, HamNoSys Version 2.0*. Hamburg, Germany, Signum Press.
- Quer, J. (2016), “Reporting with and without role shift: sign language strategies of complementation”. In: R. Pfau, M. Steinbach, A. Herrman (eds.), *A Matter of Complexity: Subordination in Sign Languages*, (pp. 204-230), Boston, De Gruyter Mouton.
- Rabiner L. R., Juang B. H. (1989), “An Introduction to Hidden Markov Models”. In: *The IEEE ASSP Magazine*, 4(1), (pp. 4-16).
- Radutzky, E. (ed.) (2001), *Dizionario bilingue elementare della Lingua dei Segni Italiana LIS*. Roma, Edizioni Kappa.
- Ramos-Soto A., Bugarin A. J., Barro S., Taboada J. (2015), “Linguistic Descriptions for Automatic Generation of Textual Short-Term Weather Forecasts on Real Prediction Data”. In: *IEEE Transactions on Fuzzy Systems*, 23(1), (pp. 44-57).
- Reape M., Mellish C. (1999), “Just what is aggregation anyway?”. In: *Proc. ENLG'99*.
- Reichenbach, H. (1947), *Elements of Symbolic Logic*, London, New York, Macmillan.
- Reilly, J. S., McIntire, M. L., Seago, H. (1992) “Affective prosody in American Sign Language”. *Sign Language Studies* 21(75), (pp 113-128).
- Reiter E. (2010), “Natural Language Generation”. In: A. Clark, C. Fox, S. Lappin (eds.), *The Handbook of Computational Linguistics and Natural Language Processing*. Wiley-Blackwell (pp. 574-598).
- Reiter E., Dale R. (1997), “Building natural-language generation systems”. In: *Natural Language Engineering*, 3, (pp. 57-87).
- Reiter E., Dale R. (2000), *Building Natural Language Generation Systems*. Cambridge, U.K., Cambridge University Press.
- Reiter E., Robertson R., Osman L., (1999), “Types of knowledge required to personalise smoking cessation letters”. In: Werner Horn *et al.* (eds.), *Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making (AIMDM'99)*, (pp. 389-399), Springer-Verlag, Berlin.
- Reiter E., Robertson R., Osman L. (2003), “Lessons from a failure: generating tailored smoking cessation letters”. In: *Artificial Intelligence 144* (pp. 41–58).

- Reiter E., Sripada S., Hunter J. R., Yu J., Davy I. (2005), “Choosing words in computer-generated weather forecasts”. In: *Artificial Intelligence, 167(1-2)*, (pp 137–169).
- Reiter E., Turner R., Alm N., Black R., Dempster M., Waller A. (2009), “Using NLG to help language-impaired users tell stories and participate in social dialogues”. In: *Proceedings of the 12th European Workshop on Natural Language Generation (ENLG-2009)*, (pp. 1-8).
- Ritchie G. (2003), *The JAPE riddle generator: technical specification*. Informatics Research Report EDI-INF-RR-0158, School of Informatics, University of Edinburgh.
- Robin J., McKeown K. R. (1996), “Empirically designing and evaluating a new revision-based model for summary generation”. In: *Artificial Intelligence 85* (pp. 135-79).
- Rus V., Wyse B., Piwek P., Lintean M., Stoyanchev S., Moldovan C. (2010). “Overview of the first question generation shared task evaluation challenge”. In: *Proc. 3rd Workshop on Question Generation*, (pp. 45-57).
- Russo Cardona T., Volterra V. (2007), *Le lingue dei segni. Storia e semiotica*. Roma, Carocci editore.
- Sáfár E., Marshall I. (2001). “The Architecture of an English-Text-to-Sign- Languages Translation System”. In: *Proceedings of the 2nd International Conference on Recent Advances in Natural Language Processing (RANLP-02)*, (pp. 223-228), Tzigov Chark, Bulgaria.
- Schembri A. (2003), “Rethinking ‘Classifiers’ in Signed Languages”. In: K. Emmorey (ed.), *Perspectives on Classifiers in Sign Language*, (pp. 3-34), Mahwah, NJ, Lawrence Erlbaum Associates.
- Schlenker P. (2017a), “Super monsters I: Attitude and Action Role Shift in sign language”. In: *Semantics and Pragmatics 10(9)*. Early access version (pp. 1-65).
- Schlenker P. (2017b), “Super monsters II: Role Shift, iconicity and quotation in sign language”. In: *Semantics and Pragmatics 10(2)*. Early access version (pp. 1-67).
- Schmidt C.A. (2016), *Handling Multimodality and Scarce Resources in Sign Language Machine Translation*, PhD Dissertation, RWTH Aachen University.

- Schroeder O. I. (1985), "A Problem in Phonological Description." In W. Stokoe and V. Volterra (eds.), *SLR '83*, Roma, Istituto di Psicologia, CNR and Silver Spring, Linstok Press.
- Sheremetyeva S., Nirenburg S., Nirenburg I. (1996), "Generating patent claims from interactive input". In: *Proceedings of the 8th International Workshop on Natural Language Generation (INLG '96)*, (pp. 61–70).
- Siddharthan A., Nenkova A., McKeown K. R. (2011), "Information Status Distinctions and Referring Expressions: An Empirical Study of References to People in News Summaries". In: *Computational Linguistics*, 37 (4), (pp. 811-842).
- Siddharthan A., (2014). "A survey of research on text simplification". In: *International Journal of Applied Linguistics*, 165(2), (pp. 259-298).
- Slobin, D. I., Hoiting, N., Anthony, M., Biederman, Y., Kuntze, M., Lindert, R., Pyers, J., Thumann, H., Weinberg, A. (2001) "Sign Language Transcription at the Level of Meaning Components: The Berkeley Transcription System (BTS)". In: *Sign Language & Linguistics*, 4 (pp. 63–96).
- Smith, R. (ed.) (2013), *HamNoSys 4.0. User guide. Draft – Version 3.0*, Institute of technology Blanchardstown.
- Stein D., Dreuw P., Ney H., Morrissey S., Way A. (2007), "Hand in Hand: Automatic Sign Language to English Translation". In: *Proceedings of the 11th Conference on Theoretical and Methodological Issues in Machine Translation (TMI-07)*, (pp. 214-220), Skövde, Sweden.
- Stein D., Schmidt C., Ney H. (2012), "Analysis, preparation, and optimization of statistical sign language machine translation". In: *Machine Translation, Vol. 26, No. 4*, (pp. 325–357).
- Stokoe, W. (1960), *Sign Language Structure: An Outline of the Visual Communication System of the American Deaf. Studies in Linguistics: Occasional Papers*. Buffalo, University of Buffalo.
- Stroppa N., Way A. (2006), "MaTrEx: DCU machine translation system for IWSLT 2006". In: *Proceedings of the 3rd International Workshop on Spoken Language Translation (IWSLT)*, (pp. 31–36). Kyoto, Japan.

- Supalla, T. (1986), “The Classifier System in American Sign Language”. In C. Craig (ed.), *Noun Classes and Categorization* (pp. 181-214), John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Supalla, T., Newport E. L. (1978), “How many seats in a chair? The derivation of nouns and verbs in American Sign Language”. In: Siple P. (ed.), *Understanding language through sign language research*, (pp 91-132). New York: Academic Press.
- Sutton, V. (1999) *Lessons in SignWriting. Textbook & Workbook*. La Jolla, CA: Deaf Action Committee for Sign Writing [2nd edition, 1st edition 1995].
- Tomaszewski, P., Farris, M. (2010), “Not by the Hands Alone: Functions of Non-Manual Features in Polish Sign Language”. In: B. Bokus (ed.), *Studies in the psychology of language and communication*. (pp. 289–320), Warszawa, Matrix.
- Thomason J., Venugopalan S., Guadarrama S., Saenko K., Mooney R. J. (2014). “Integrating Language and Vision to Generate Natural Language Descriptions of Videos in the Wild”. In: *Proc. COLING’14*, (pp. 1218-1227).
- Turner R., Sripada S., Reiter E., Davy I. (2008), “Selecting the Content of Textual Descriptions of Geographically Located Events in Spatio-Temporal Weather Data”. In: *Applications and Innovations in Intelligent Systems XV*, (pp. 75–88).
- Valli C., Ceil L., (2000). *Linguistics of American Sign Language: An Introduction*. Washington, D.C.: Gallaudet University Press.
- van Deemter K, Odijk J. (1997), “Context modeling and the generation of spoken discourse”. In: *Speech Communication*, 21 (pp. 101-121).
- Veale T., Conway A. (1994), “Cross Modal Comprehension in ZARDOZ an English to Sign-language Translation System”. In: *Proceedings of the Seventh International Workshop on Natural Language Generation*, (pp. 249–252). Association for Computational Linguistics, Stroudsburg, PA, USA.
- Veale T., Conway A., Collins B. (1998), “The Challenges of Cross-Modal Translation: English to Sign Language Translation in the Zardozy System”. In: *Machine Translation*, 13(1), (pp. 81–106).
- Veale T., Cunningham P. (1992), “Competitive Hypothesis Resolution in TWIG: A Blackboard-Driven Text-Understanding System”. In: *10th European Conference on Artificial Intelligence, Vienna, Austria*, (pp. 105–117).

- Verdirosi, M. L. (1987), “Luoghi”. In V. Volterra (ed.), *La Lingua Italiana dei Segni La comunicazione visivo-gestuale dei sordi*. (pp 23-48). Bologna: Il Mulino. (Nuova Edizione 2004).
- Vogt-Svendsen, M. (1984). “Word-Pictures in Norwegian Sign Language (NSL). A preliminary Analysis”. In: *Working Papers in Linguistics 2*, (pp 112- 141). University of Trondheim.
- Volterra, V. (1987), *La Lingua Italiana dei Segni La comunicazione visivo-gestuale dei sordi*. Bologna: Il Mulino. (Nuova Edizione 2004).
- Volterra, V., Laudanna, A., Corazza, S., Radutzky, E., Natale, F. (1984), “Italian Sign Language: the order of elements in the declarative sentence”. In: F. Loncke, P. Boyes-Braem, Y. Lebrun (eds.), *Recent research on European Sign Languages*, (pp. 19-48), Lisse, Swets & Zeitlinger.
- White M., Rajkumar R. (2012), “Minimal dependency length in realization ranking”. In: *Proc. EMNLP’12*, (pp. 244-255), Jeju Island, Korea.
- Wilbur, R. B. (2009), “Phonological and Prosodic Layering of Nonmanuals in American Sign Language”. In: K. Emmorey, H. Lane (eds.), *The Signs of Language Revisited: an anthology to honor Ursula Bellugi and Edward Klima*, (pp 215-244). Mahwah, NJ, Lawrence Erlbaum Associates.
- Worsley, P. M. (1954), “Noun classification in Australian and Bantu: formal or semantic?”. In: *Oceania*, 24, (pp. 275-288).
- Wu C.-H., Su H.-Y., Chiu Y.-H., and Lin C.-H. (2007), “Transfer-Based Statistical Translation of Taiwanese Sign Language Using PCFG”. *ACM Transactions on Asian Language Information Processing (TALIP)*, 6(1).
- Yu J., Reiter E., Hunter J., Mellish C. (2007), “Choosing the content of textual summaries of large time-series data sets”. In: *Natural Language Engineering 13*, (pp. 25-49).
- Zhao L., Kipper K., Schuler W., Vogler C., Badler N., Palmer M. (2000), “A Machine Translation System from English to American Sign Language”. In: *Envisioning Machine Translation in the Information Future: Proceedings of the Fourth Conference of the Association for Machine Translation (AMTA-00)*, (pp. 293-300), Cuernavaca, Mexico.

Zucchi, S. (2004), *Monsters in the visual mode?* Milano, Università degli Studi di Milano.

Zucchi, S. (2009), “Along the Time Line: Tense and Time Adverbs in Italian Sign Language”, (pp. 99-139). In: *Natural Language Semantics*, 17.

Video: “La lepre e la tartaruga”, in: *Fiabe nel Bosco I*. DVD, Alba Cooperativa Sociale ONLUS, 2010.

## Webography

*Creating StorySign - an app for deaf readers* – last consulted 4/08/19

<[aardman.com/blog/creating-storysign-app-for-deaf-readers](http://aardman.com/blog/creating-storysign-app-for-deaf-readers)>

*Discover the Magic of StorySign* – last consulted 4/08/19

<[consumer.huawei.com/en/campaign/storysign](http://consumer.huawei.com/en/campaign/storysign)>

*Downloadable NLG systems* – last consulted 4/08/19

<[aclweb.org/aclwiki/Downloadable\\_NLG\\_systems](http://aclweb.org/aclwiki/Downloadable_NLG_systems)>

*ELAN - The Language Archive* – last consulted 19/07/19

<[tla.mpi.nl/tools/tla-tools/elan](http://tla.mpi.nl/tools/tla-tools/elan)>

*Ethnologue: Languages of the World* – last consulted 9/07/19

<[ethnologue.com](http://ethnologue.com)>

*HUAWEI StorySign App - Available Now* – last consulted 4/08/19

<[youtube.com/watch?v=rQazllUknIU](https://youtube.com/watch?v=rQazllUknIU)>

*ItalWordNet* – last consulted 10/09/19

<[cnr.it/it/banche-dati-istituti/banca-dati/442/italwordnet-iwn](http://cnr.it/it/banche-dati-istituti/banca-dati/442/italwordnet-iwn)>

*Kinect Sign Language Translator - part 1* – last consulted 10/08/19

<[microsoft.com/en-us/research/blog/kinect-sign-language-translator-part-1](http://microsoft.com/en-us/research/blog/kinect-sign-language-translator-part-1)>

*MT talks - Intro* – last consulted 07/08/19

<[mttalks.ufal.ms.mff.cuni.cz/index.php?title=Intro](http://mttalks.ufal.ms.mff.cuni.cz/index.php?title=Intro)>

*Scientific understanding and vision-based technological development for continuous sign language recognition and translation* – last consulted 9/08/19

<[signspeak.eu/en/overall.html](http://signspeak.eu/en/overall.html)>

*SignAll - Real Time Automated ASL to English Translation Technology* – last consulted 12/08/19

<[signall.us](http://signall.us)>

*SignAll @Gallaudet Uni 2018* – last consulted 12/08/19

<[youtube.com/watch?v=hx1kn55TWpY](https://youtube.com/watch?v=hx1kn55TWpY)>

*SignAll's Learn ASL Introduction* – last consulted 12/08/19

<[youtube.com/watch?v=IE5bBLk51hw](https://www.youtube.com/watch?v=IE5bBLk51hw)>

*SignSpeak - Project description* – last consulted 9/08/19

<[ru.nl/sign-lang/projects/completed-projects/signspeak](http://ru.nl/sign-lang/projects/completed-projects/signspeak)>

*Spreadthesign* – last consulted 10/07/19

<[spreadthesign.com](http://spreadthesign.com)>