Workshop Proceedings

6th Workshop on the Representation and Processing of Sign Languages:
Beyond the Manual Channel

Language Resources and Evaluation Conference (LREC)
Reykjavik, Iceland, 31 May 2014
Editors and Workshop Organizers

Onno Crasborn  
Radboud University, Nijmegen NL

Eleni Efthimiou  
Institute for Language and Speech Processing, Athens GR

Evita Fotinea  
Institute for Language and Speech Processing, Athens GR

Thomas Hanke  
Institute of German Sign Language, University of Hamburg, Hamburg DE

Julie Hochgesang  
Gallaudet University, Washington US

Jette Kristoffersen  
Centre for Sign Language, University College Capital, Copenhagen DK

Johanna Mesch  
Stockholm University, Stockholm SE

Workshop Programme Committee

Richard Bowden  
University of Surrey, Guildford GB

Penny Boyes Braem  
Center for Sign Language Research, Basel CH

Annelies Braffort  
LIMSI/CNRS, Orsay FR

Christophe Collet  
IRIT, University of Toulouse, Toulouse FR

Kearsy Cormier  
Deafness Cognition and Language Research Centre, London GB

Onno Crasborn  
Radboud University, Nijmegen NL

Svetlana Dachkovsky  
University of Haifa, Haifa IL

Eleni Efthimiou  
Institute for Language and Speech Processing, Athens GR

Stavroula-Evita Fotinea  
Institute for Language and Speech Processing, Athens GR

John Glauert  
University of East Anglia, Norwich GB

Thomas Hanke  
Institute of German Sign Language, University of Hamburg, Hamburg DE

Alexis Heloir  
German Research Centre for Artificial Intelligence, Saarbrücken DE

Jens Heßmann  
University of Applied Sciences Magdeburg-Stendal, Magdeburg DE

Julie Hochgesang  
Gallaudet University, Washington US

Trevor Johnston  
Macquarie University, Sydney AU

Reiner Konrad  
Institute of German Sign Language, University of Hamburg, Hamburg DE

Jette Kristoffersen  
Centre for Sign Language, University College Capital, Copenhagen DK

Lorraine Leeson  
Trinity College, Dublin IE

Petros Maragos  
National Technical University of Athens, Athens GR

John McDonald  
DePaul University, Chicago US

Johanna Mesch  
Stockholm University, Stockholm SE

Carol Neidle  
Boston University, Boston US

Christian Rathmann  
Institute of German Sign Language, University of Hamburg, Hamburg DE

Adam Schembri  
National Institute for Deaf Studies and Sign Language, La Trobe University, Melbourne AU

Rosalee Wolfe  
DePaul University, Chicago US
# Table of contents

Mouth features as non-manual cues for the categorization of lexical and productive signs in French Sign Language (LSF)  
Antonio Balvet and Marie-Anne Sallandre  
1

Segmenting the Swedish Sign Language corpus: On the possibilities of using visual cues as a basis for syntactic segmentation  
Carl Börstell, Johanna Mesch and Lars Wallin  
7

Synthesizing facial expressions for sign language avatars  
Yosra Bouzid, Oussama El Ghoul and Mohamed Jemni  
11

Eye gaze annotation practices: Description vs. interpretation  
Annelies Braffort  
19

An annotation scheme for mouth actions in sign languages  
Onno Crasborn and Richard Bank  
23

Implementation of an automatic sign language lexical annotation framework based on Propositional Dynamic Logic  
Arturo Curiel and Christophe Collet  
29

Creation of a multipurpose sign language lexical resource: The GSL lexicon database  
Athanasia-Lida Dimou, Theodore Goulas, Eleni Efthimiou, Stavroula-Evita Fotinea, Panagiotis Karioris, Michalis Pissaris, Dimitis Korakakis and Kiki Vasilaki  
37

A hybrid formalism to parse sign languages  
Rémi Dubot and Christophe Collet  
43

Non-manual features: The right to indifference  
Michael Filhol, Mohamed Nassime Hadjadj and Annick Choisier  
49

When nonmanuals meet semantics and syntax: Towards a practical guide for the segmentation of sign language discourse  
Sílvia Gabarró-López and Laurence Meurant  
55

Last train to “Rebaudengo Fossano”: The case of some names in avatar translation  
Carlo Geraci and Alessandro Mazzei  
63

Annotation of mouth activities with iLex  
Thomas Hanke  
67

Release of experimental stimuli and questions for evaluating facial expressions in animations of American Sign Language  
Matt Huenerfauth and Hernisa Kacorri  
71

How to use depth sensors in sign language corpus recordings  
Rekha Jayaprakash and Thomas Hanke  
77

Mouth-based non-manual coding schema used in the Auslan corpus: Explanation, application and preliminary results  
Trevor Johnston and Jane van Roekel  
81

Weakly supervised automatic transcription of mouthings for gloss-based sign language corpora  
Oscar Koller, Hermann Ney and Richard Bowden  
89
Discourse-based annotation of relative clause constructions in Turkish Sign Language (TID): A case study
Okan Kubus

Signing thoughts! A methodological approach within the semantic field work used for coding nonmanuals which express modality in Austrian Sign Language (ÖGS)
Andrea Lackner and Nikolaus Riemer

Estimating head pose and state of facial elements for sign language video
Marcos Luzardo, Ville Viitaniemi, Matti Karppa, Jorma Laaksonen and Tommi Jantunen

Addressing the cardinals puzzle: New insights from non-nanual markers in Italian Sign Language
Lara Mantovan, Carlo Geraci and Anna Cardinaletti

Analysis for synthesis: Investigating corpora for supporting the automatic generation of role shift
John McDonald, Rosalee Wolfe, Robyn Moncrief and Souad Baowidan

The “how-to” of integrating FACS and ELAN for analysis of non-manual features in ASL
Kristin Mulrooney, Julie Hochgesang, Carla Morris and Katie Lee

Computer-based tracking, analysis, and visualization of linguistically significant nonmanual events in American Sign Language (ASL)
Carol Neidle, Jingjing Liu, Bo Liu, Xi Peng, Christian Vogler and Dimitris Metaxas

Nonmanuals and markers of (dis)fluency
Ingrid Notarrigo and Laurence Meurant

Taking non-manuality into account in collecting and analyzing Finnish Sign Language video data
Anna Puupponen, Tommi Jantunen, Ritva Takkinen, Tuia Wainio and Outi Pippuri

Visualizing the spatial working memory in mathematical discourse in Finnish Sign Language
Päivi Rainö, Marja Huovila and Irja Seilola

Use of nonmanuals by adult L2 signers in Swedish Sign Language – Annotating the nonmanuals
Krister Schönström and Johanna Mesch
## Author Index

Balvet, Antonio  
Bank, Richard  
Baowidan, Souad  
Bouzid, Yosra  
Bowden, Richard  
Börstell, Carl  
Braffort, Annelies  
Cardinaletti, Anna  
Choisier, Annick  
Collet, Christophe  
Crasborn, Onno  
Curiel, Arturo  
Dimou, Athanasia-Lida  
Dubot, Rémi  
Efthimiou, Eleni  
El Ghoul, Oussama  
Filhol, Michael  
Fotinea, Stavroula-Evita  
Gabarró-López, Silvia  
Geraci, Carlo  
Goulas, Theodore  
Hanke, Thomas  
Hochgesang, Julie  
Huenerfauth, Matt  
Huovila, Marja  
Jantunen, Tommi  
Jayaprakash, Rekha  
Jenni, Mohamed  
Johnston, Trevor  
Kacorri, Hernisa  
Karioris, Panagiotis  
Karppa, Matti  
Koller, Oscar  
Korakakis, Dimitis  
Kubus, Okan  
Laaksonen, Jorma  
Lackner, Andrea  
Lee, Katie  
Liu, Bo  
Liu, Jingjing  
Luzardo, Marcos  
Mantovan, Lara  
Mazzei, Alessandro  
McDonald, John  
Mesch, Johanna  
Metaxas, Dimitris  
Meurant, Laurence  
Moncrief, Robyn  
Morris, Carla

1, 23, 117, 11, 89, 7, 19, 117, 49, 49, 7, 19, 37, 43, 37, 37, 37, 49, 37, 55, 63, 113, 7, 153, 127, 127, 105, 105, 113, 63, 117, 7, 153, 127, 55, 135, 117, 123
<table>
<thead>
<tr>
<th>Name</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mulrooney, Kristin</td>
<td>123</td>
</tr>
<tr>
<td>Nassime, Mohamed</td>
<td>49</td>
</tr>
<tr>
<td>Neidle, Carol</td>
<td>127</td>
</tr>
<tr>
<td>Ney, Hermann</td>
<td>89</td>
</tr>
<tr>
<td>Notarrigo, Ingrid</td>
<td>135</td>
</tr>
<tr>
<td>Peng, Xi</td>
<td>127</td>
</tr>
<tr>
<td>Pippuri, Outi</td>
<td>143</td>
</tr>
<tr>
<td>Pissaris, Michalis</td>
<td>37</td>
</tr>
<tr>
<td>Puupponen, Anna</td>
<td>143</td>
</tr>
<tr>
<td>Rainò, Päivi</td>
<td>149</td>
</tr>
<tr>
<td>Riemer, Nikolaus</td>
<td>100</td>
</tr>
<tr>
<td>Sallandre, Marie-Anne</td>
<td>1</td>
</tr>
<tr>
<td>Schönström, Krister</td>
<td>153</td>
</tr>
<tr>
<td>Seilola, Irja</td>
<td>149</td>
</tr>
<tr>
<td>Takkinen, Ritva</td>
<td>143</td>
</tr>
<tr>
<td>van Roekel, Jane</td>
<td>81</td>
</tr>
<tr>
<td>Vasilaki, Kiki</td>
<td>37</td>
</tr>
<tr>
<td>Viitaniemi, Ville</td>
<td>105</td>
</tr>
<tr>
<td>Vogler, Christian</td>
<td>127</td>
</tr>
<tr>
<td>Wainio, Tuija</td>
<td>143</td>
</tr>
<tr>
<td>Wallin, Lars</td>
<td>7</td>
</tr>
<tr>
<td>Wolfe, Rosalee</td>
<td>117</td>
</tr>
</tbody>
</table>
Editors’ Preface

This collection of papers stems from the Sixth Workshop on the Representation and Processing of Sign Languages, held in May 2014 as a satellite to the Language Resources and Evaluation Conference in Reykjavik. While there has been occasional attention for sign languages at the main LREC conference, the main focus there is on spoken languages in their written and spoken forms. This series of workshops, however, offers a forum for researchers focusing on sign languages. For the fourth time, the workshop had sign language corpora as its main topic. This time, the focus was on any aspect beyond the manual channel. Not surprisingly, most papers deal with non-manuals on the face.

Once again, the papers at this workshop clearly identify the potentials of even closer cooperation between sign linguists and sign language engineers, and we think it is events like this that contribute a lot to a better understanding between researchers with completely different backgrounds.

The contributions composing this volume are presented in alphabetical order by the first author. For the reader’s convenience, an author index is provided as well.

We would like to thank all members of the programme committee who helped us reviewing the submissions to the workshop within a very short timeframe!

Finally, we would like to point the reader to the proceedings of the previous workshops that form important resources in a growing field of research:


The Editors
Addressing the Cardinals Puzzle:
New Insights from Non-Manual Markers in Italian Sign Language

Lara Mantovan¹, Carlo Geraci², Anna Cardinaletti¹
Ca’ Foscari University of Venice¹, Institut Jean-Nicod CNRS²

Address: Lara Mantovan, Dorsoduro 1075, Fondamenta Tofetti, 30123 Venezia
Email: laramantovan@unive.it, carlo.geraci76@gmail.com, cardin@unive.it

Abstract

This paper aims at investigating the main linguistic properties associated with cardinal numerals in LIS (Italian sign language). Considering this issue from several perspectives (phonology, prosody, semantics and syntax), we discuss some relevant corpus and elicited data with the purpose of shedding light on the distribution of cardinals in LIS. We also explain what triggers the emergence of different word/sign orders in the noun phrase. Non-manual markers are crucial in detecting two particular subcases.

Keywords: cardinal numerals, nonmanuals, Italian sign language, noun phrases, sign order

1. Background

In this paper we focus on cardinal numerals functioning as modifiers in the nominal domain and expressing a certain quantity. The cardinal system in Italian sign language (LIS) uses both hands and is a base-10 system. In this respect, the distribution of cardinals in LIS reveals a puzzling picture. On the one hand, recent corpus data from 162 LIS signers reveal that in spontaneous narratives the majority of cardinals appears before the noun (Mantovan & Geraci, 2013), as reported in Table 1.

<table>
<thead>
<tr>
<th>Word order</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Card &gt; N</td>
<td>278</td>
<td>79%</td>
</tr>
<tr>
<td>N &gt; Card</td>
<td>75</td>
<td>21%</td>
</tr>
</tbody>
</table>

Table 1: Distribution of cardinal numerals in corpus data

On the other hand, the existing literature claims that cardinals are consistently or even exclusively postnominal (Bertone, 2007; Branchini, 2007; Cecchetto, Geraci & Zucchi, 2009; Brunelli, 2011). An example from Bertone (2007) is reported below for expository purposes.

[Bertone, 2007:84]

(1) BOOK NEW TWO DEM MINE
‘These two new books are mine.’

Why do we observe such an important difference between corpus data and elicited data? In what respect is Card>N different from N>Card (and vice versa)? In the remainder of the paper we will offer an explanation for these two newly discovered puzzles. Our working hypothesis is that part of the sign order variability is due to the definite/indefinite character of the noun phrase, that is marked both by prosodic (i.e. non-manual) features and sign order manipulation.

2. Methods

The data for this study mainly come from the LIS corpus (Geraci et al., 2011). The annotated cardinals amount to 353 tokens. Additional data have been collected through picture-based narration tasks and elicitation of grammaticality judgments.

The materials used as stimuli for the picture-based narration tasks are wordless comic strips illustrated by Plauen (2000). Plauen’s illustrations are generally self-explanatory and do not give rise to interlinguistic influences since they do not contain any written text.

Figure 1: Extract from “Vater und Sohn” (Plauen 2000)

The story represented in Figure 1 is interesting because it triggers the production of cardinal TWO in two different contexts. In the first panel two children are represented for the very first time. Being first-mentioned referents, they are expected to be introduced in the discourse by an
indefinite noun phrase. On the contrary, the two children represented in the fifth panel are pre-established referents, therefore they are expected to be referred to by using a definite noun phrase.

Data annotation has been conducted by using the annotation software ELAN (Johnston & Crasborn, 2006). Manual and non-manual features have been carefully annotated on separate tiers. The coding scheme associated to the non-manual markers (NMMs) relevant for this study is illustrated in (2). The duration of NMMs has been measured as the time interval intervening between start and end points.

(2) a. NM-Head: left, right, raised, down, forward, back
b. NM-Eyebrows: lowered, raised
d. NM-Body: left, right, down, forward, back
e. NM-Eyes: blink, squint, close, wide, track-hands eye-gaze

To illustrate how ELAN has been used for data annotation, a representative screenshot is shown in Figure 2.

Figure 2: ELAN screenshot

Finally, grammaticality judgments have been elicited from three native signers of LIS (Rosella Ottolini, Gabriele Caia and Mirko Santoro), whom we thank enormously.

3. Results

A deeper investigation of the distribution of cardinals as emerging from the LIS corpus (see Table 1 above) revealed the presence of a confounder, namely the potentially ambiguous status of the sign ONE, and the special behavior of a subclass of cardinals, namely the ones contained in measure phrases. We discuss each of them in turn.

3.1. The sign ONE

Similarly to “uno/una” in Italian, the LIS sign ONE is ambiguous between a cardinal and an indefinite determiner. In our corpus, ONE mainly occurs in prenominal position (almost 90% of the cases) irrespectively of the syntactic/semantic function. The distribution of determiner ONE and cardinal ONE can be observed in examples (3) and (4), respectively (see also Figure 3 and Figure 4).

Corpus data (middle-aged signer from Rome)

(3) ONE MATE SCHOOL IX-3 POSS IX-3 JEALOUS STRONG
'A schoolmate of mine was extremely jealous.'

Corpus data (middle-aged signer from Rome)

(4) REFECTORY EAT FINISHED, REFECTORY ARRANGE TURN, ONE WEEK IX-1, THEN WEEK IX-3
'After we finished eating at the refectory, we took turns arranging things, one week it was my turn, then it was someone else's turn.'

As originally suggested by Bertone (2007), NMMs help distinguish the two functions. Figure 3 shows the facial expressions associated with determiner ONE in sentence (3). The most remarkable features are backward-tilted head and raised eyebrows.

Figure 3: ONE as indefinite determiner

Figure 4 shows the realization of cardinal ONE in sentence (4). In this latter case, eyebrows are in neutral position and the head is not backward tilted.

Figure 4: ONE as cardinal numeral

Once the occurrences of ONE are removed from the counting, we obtain the distribution represented in Table 2.
3.2. Cardinals within Measure Phrases

Let’s now turn to the special case of cardinals included in measure phrases referring to time, capacity, weight, length, temperature, currency (e.g. SIX WEEK, SEVENTY KILOGRAM, THIRTY KILOMETER). According to corpus data, they display a categorical distribution: they always precede the noun, as shown in examples in (5) and (6).

Corpus data (young signer from Lamezia)
(5) NOW IX-3_POSS WIFE PREGNANT FIVE MONTH ‘Now my wife is five months pregnant.’

Corpus data (old signer from Florence)
(6) HOUSE NEAR, FOUR-HUNDRED METER IX-3 ‘The house is in the neighborhood, about four hundred meters away.’

This piece of data has been confirmed by grammaticality judgments, as exemplified in (7).

(7) a. IX-1 REPEAT++ TWO-HUNDRED-THOUSAND TIME ‘I repeated it two hundred thousand times.’
   b. * IX-1 REPEAT++ TIME TWO-HUNDRED-THOUSAND

Without considering these two special cases, the distribution of cardinals, shown in Table 3, looks considerably different from the ones reported in Table 1 and Table 2. As a result, the percentage of postnominal cardinals becomes more prominent and it is now perfectly balanced with prenominal cardinals.

<table>
<thead>
<tr>
<th>Word order</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Card &gt; N</td>
<td>184/252</td>
<td>73%</td>
</tr>
<tr>
<td>N &gt; Card</td>
<td>68/252</td>
<td>27%</td>
</tr>
</tbody>
</table>

Table 2: Distribution of cardinal numerals without ONE

Table 3: Distribution of cardinal numerals without ONE and measure phrases

The picture that emerges is even more intricate, showing an apparently uncontrolled variability. We now turn to the narration tasks and grammaticality judgment elicitation in order to address this issue.

3.3. The distribution of cardinals

The data collected during the narration tasks and elicitation reveal that the position of cardinals may be influenced by information structure. New-discourse information (e.g. first-mentioned referents) can be conveyed by both orders (i.e. Card>N and N>Card), whereas old-discourse information (i.e. already-mentioned referents) is compatible with N>Card only. The former is illustrated in the first panel of the comic strip, shown here in Figure 5; the latter in the fifth panel, shown here in Figure 6.

Figure 5: First-mentioned referents (new-discourse information)

Figure 6: Already-mentioned referents (old-discourse information)

When the children are first mentioned we observe both orders Card>N and N>Card, while in further mentioning only the N>Card order is found.

This is further confirmed by the informants’ assessment of their own productions. When explicitly asked about the order possibility in the two distinct contexts, only the new-information situation allows for the two sign order options, as exemplified in (8). On the contrary, in the old-discourse context only the N>Card order is possible, as illustrated in (9).

(8) New-information context
   a. TWO CHILD
   b. CHILD TWO
   ‘Two children’
Old-information context
a. * TWO CHILD
b. CHILD TWO
'The two children'

It is worth noting that the relative order of cardinals with respect to the noun is not crucial to distinguish the two discourse functions, as the sequentially identical data in (8)b and (9)b demonstrate. Rather, we found that it is the NMM component that plays a crucial role here. If the signer is dealing with a new referent, the prenominal or postnominal cardinal is usually accompanied by backward-tilted head and raised eyebrows (see Figure 7).

If the referent has already been mentioned in the discourse, then the postnominal cardinal is compatible with squinted eyes and/or lowered eyebrows (see Figure 8).

We tentatively associate the new/old discourse information with the [+definite] character of the noun phrase. Interestingly, when the noun phrase is new information, it is introduced by the same NMMs as indefinite ONE and the prenominal syntactic position is available for cardinals.

From a syntactic point of view, in the spirit of Cardinaletti and Giusti (2006), the former cardinal functions as a proper quantifier, whereas the latter, being compatible with a definite environment, should be rather considered as a quantity adjective.

4. Conclusions

In this study we combined both quantitative and qualitative data with the purpose of capitalizing on the advantages of each source. When analyzing cardinals in LIS, two special cases (i.e. ONE and cardinals within measure phrases) need to be examined separately. Syntactic positions and, most importantly, NMMs convey crucial information on the definite or indefinite nature of the nominal expression containing cardinal numerals.

5. References


