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 Radboud University, Nijmegen NL Onno Crasborn 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 Gallaudet University, Washington US Julie Hochgesang **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 Trinity College, Dublin IE Lorraine Leeson National Technical University of Athens, Athens GR Petros Maragos John McDonald DePaul University, Chicago US Johanna Mesch Stockholm University, Stockholm SE Boston University, Boston US Carol Neidle Christian Rathmann Institute of German Sign Language, University of Hamburg, Hamburg DE National Institute for Deaf Studies and Sign Language, La Adam Schembri 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) <i>Antonio Balvet and Marie-Anne Sallandre</i>	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 <i>Onno Crasborn and Richard Bank</i>	23
Implementation of an automatic sign language lexical annotation framework based on Propositional Dynamic Logic <i>Arturo Curiel and Christophe Collet</i>	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 <i>Rémi Dubot and Christophe Collet</i>	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 <i>Sílvia Gabarró-López and Laurence Meurant</i>	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 <i>Matt Huenerfauth and Hernisa Kacorri</i>	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 <i>Trevor Johnston and Jane van Roekel</i>	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	95
Signing thoughts! A methodological approach within the semantic field work used for coding nonmanuals which express modality in Austrian Sign Language (ÖGS) <i>Andrea Lackner and Nikolaus Riemer</i>	100
Estimating head pose and state of facial elements for sign language video Marcos Luzardo, Ville Viitaniemi, Matti Karppa, Jorma Laaksonen and Tommi Jantunen	105
Addressing the cardinals puzzle: New insights from non-nanual markers in Italian Sign Language	112
Analysis for synthesis: Investigating corpora for supporting the automatic generation of	115
role shift John McDonald, Rosalee Wolfe, Robyn Moncrief and Souad Baowidan	117
The "how-to" of integrating FACS and ELAN for analysis of non-manual features in ASL <i>Kristin Mulrooney, Julie Hochgesang, Carla Morris and Katie Lee</i>	123
Computer-based tracking, analysis, and visualization of linguistically significant nonmanual events in American Sign Language (ASL) Carol Neidle Jingiing Liu, Bo Liu, Xi Peng, Christian Vogler and Dimitris Metaxas	127
Nonmanuals and markers of (dis)fluency Ingrid Notarrigo and Laurence Meurant	135
Taking non-manuality into account in collecting and analyzing Finnish Sign Language video data	
Anna Puupponen, Tommi Jantunen, Ritva Takkinen, Tuija Wainio and Outi Pippuri	143
Visualizing the spatial working memory in mathematical discourse in Finnish Sign Language	
Päivi Rainò, Marja Huovila and Irja Seilola	149
Use of nonmanuals by adult L2 signers in Swedish Sign Language – Annotating the nonmanuals	
Krister Schönström and Johanna Mesch	153

Author Index

Balvet, Antonio	1
Bank, Richard	23
Baowidan, Souad	117
Bouzid, Yosra	11
Bowden, Richard	89
Börstell, Carl	7
Braffort, Annelies	19
Cardinaletti, Anna	113
Choisier, Annick	49
Collet, Christophe	29, 43
Crasborn, Onno	23
Curiel, Arturo	29
Dimou, Athanasia-Lida	37
Dubot, Rémi	43
Efthimiou, Eleni	37
El Ghoul, Oussama	11
Filhol, Michael	49
Fotinea, Stavroula-Evita	37
Gabarró-López, Sílvia	55
Geraci, Carlo	63, 113
Goulas, Theodore	37
Hanke, Thomas	67, 77
Hochgesang, Julie	123
Huenerfauth, Matt	71
Huovila, Marja	149
Jantunen, Tommi	105, 143
Jayaprakash, Rekha	77
Jemni, Mohamed	11
Johnston, Trevor	81
Kacorri, Hernisa	71
Karioris, Panagiotis	37
Karppa, Matti	105
Koller, Oscar	89
Korakakis, Dimitis	37
Kubus, Okan	95
Laaksonen, Jorma	105
Lackner, Andrea	100
Lee, Katie	123
Liu, Bo	127
Liu, Jingjing	127
Luzardo, Marcos	105
Mantovan, Lara	113
Mazzei, Alessandro	63
McDonald, John	117
Mesch, Johanna	7, 153
Metaxas, Dimitris	127
Meurant, Laurence	55, 135
Moncrief, Robyn	117
Morris, Carla	123

Mulrooney, Kristin	123
Nassime, Mohamed	49
Neidle, Carol	127
Ney, Hermann	89
Notarrigo, Ingrid	135
Peng, Xi	127
Pippuri, Outi	143
Pissaris, Michalis	37
Puupponen, Anna	143
Rainò, Päivi	149
Riemer, Nikolaus	100
Sallandre, Marie-Anne	1
Schönström, Krister	153
Seilola, Irja	149
Takkinen, Ritva	143
van Roekel, Jane	81
Vasilaki, Kiki	37
Viitaniemi, Ville	105
Vogler, Christian	127
Wainio, Tuija	143
Wallin, Lars	7
Wolfe, Rosalee	117

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 focussing 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:

- O. Streiter & C. Vettori (2004, Eds.) From SignWriting to Image Processing. Information techniques and their implications for teaching, documentation and communication. [Proceedings of the Workshop on the Representation and Processing of Sign Languages. 4th International Conference on Language Resources and Evaluation, LREC 2004, Lisbon.] Paris: ELRA. Available online at <u>http://www.lrecconf.org/proceedings/lrec2004/ws/ws18.pdf</u>
- C. Vettori (2006, Ed.) *Lexicographic Matters and Didactic Scenarios*. [Proceedings of the 2nd Workshop on the Representation and Processing of Sign Languages. 5th International Conference on Language Resources and Evaluation, LREC 2006, Genova.] Paris: ELRA. Available online at <u>http://www.lrec-conf.org/proceedings/lrec2006/workshops/W15/Sign_Language_Workshop_Proceedings.pdf</u>
- O. Crasborn, E. Efthimiou, T. Hanke, E. Thoutenhoofd & I. Zwitserlood (2008, Eds.) Construction and Exploitation of Sign Language Corpora. [Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages. 6th International Conference on Language Resources and Evaluation, LREC 2008, Marrakech.] Paris: ELRA. Available online at <u>http://www.lrecconf.org/proceedings/lrec2008/workshops/W25_Proceedings.pdf</u>
- P. Dreuw, E. Efthimiou, T. Hanke, T. Johnston, G. Martínez Ruiz & A. Schembri (2010, Eds.) *Corpora and Sign Language Technologies*. [Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages. 7th International Conference on Language Resources and Evaluation, LREC 2010, Valletta, Malta.] Paris: ELRA. Available online at <u>http://www.lrec-conf.org/proceedings/lrec2010/workshops/W13.pdf</u>
- O. Crasborn, E. Efthimiou, S.-E. Fotinea, T. Hanke, J. Kristoffersen, J. Mesch (2012, Eds.) *Interaction between Corpus and Lexicon*. [Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages. 8th International Conference on Language Resources and Evaluation, LREC 2012, Istanbul, Turkey.] Paris: ELRA. Available online at http://www.lrec-conf.org/proceedings/lrec2012/workshops/24.Proceedings_SignLanguage.pdf

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.

Word order	n	%
Card > N	278/353	79%
N > Card	75/353	21%

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 selfexplanatory 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.

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)

 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.

Word order	n	%
Card > N	184/252	73%
N > Card	68/252	27%

Table 2: Distribution of cardinal numerals without ONE

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)

 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.

Word Order	n	%
Card > N	67 /135	50%
N > Card	68 /135	50%

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 olddiscourse context only the N>Card order is possible, as illustrated in (9).

(8) New-information context
 a. TWO CHILD
 b. CHILD TWO
 'Two children'

(9) 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).



Figure 7: TWO as new-discourse information

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).



Figure 8: TWO as old-discourse information

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

- Bertone, C. (2007). La struttura del sintagma determinante nella Lingua dei Segni Italiana (LIS). Ph.D. Thesis in Linguistics. University Ca' Foscari of Venice.
- Branchini, C. (2007). On Relativization and Clefting in Italian Sign Language (LIS), Ph.D. Thesis in Linguistics. University Ca' Foscari of Venice.
- Brunelli, M. (2011). Antisymmetry and Sign Languages. A comparison between NGT and LIS, Utrecht: LOT Publications.
- Cardinaletti, A. & Giusti, G. (2006). The syntax of quantified phrases and quantitative clitics. In M. Everaert, H. van Riemsdijk (eds.), The Blackwell Companion to Syntax, Oxford: Blackwell, vol. 5, pp. 23-93.
- Cecchetto, C., Geraci, C. & Zucchi, A. (2009). Another way to mark syntactic dependencies: The case for rightperipheral specifiers in sign languages. In Language, 85:2, pp. 278-320.
- Geraci, C., Battaglia, K., Cardinaletti, A., Cecchetto, C., Donati, C., Giudice, S. & Mereghetti, E. (2011). The LIS Corpus Project. A Discussion of Sociolinguistic Variation in the Lexicon. In Sign Language Studies, 11:4, pp. 528-574.
- Johnston, T. & Crasborn, O. (2006). The use of ELAN software annotation software in the creation of sign language corpora. Presentation at the E-MELD workshop on digital language documentation. Michigan State University in East Lansing, Michigan, US, June 2006.
- Mantovan, L. & Geraci, C. (2013). A round trip from Theory to Corpus. The Case of Universal 20 in LIS. Poster presented at the TISLR 11 Conference, London, July 2013.
- Pfau, R. & Quer, J. (2010). Nonmanuals: their grammatical and prosodic roles. In D. Brentari (ed.), Sign Languages, Cambridge: Cambridge University Press, pp. 381-402.
- Plauen, E. O. (2000). Vater und Sohn. In Politische Karikaturen, Zeichnungen, Illustrationen und alle Bildgeschichten "Vater und Sohn" (Gesamtausgabe), Konstanz: Südverlag GmbH.