Keeping track in sign language: coreference annotation using Centering Theory

Gemma Barberà & Guillem Massó
Universitat Pompeu Fabra

___Sign Language Corpora : Linguistic Issues___
University College London – 24 July 2009
Main goals

• To establish a possible way of coreference annotation in SLs
• To describe a hierarchy of referring expressions in LSC (Catalan Sign Language)
• To start identifying the linguistic constraints that characterize coreferential expressions
Main claims

• Referring expressions can be distinguished as to the degree in which their referents are accessible at different points in a discourse

• The processing effort to process a piece of discourse varies and this is reflected in the choice of referring expressions

• Theoretical approaches predictions:
  • nonsalient or distant antecedents → anaphora coded by a lower accessibility marker (more informative)
  • salient or recently mentioned antecedent → high accessible marker (less informative)
Discourse research in OLs: Ranking of saliency

• Prince (1981): provides a taxonomy of different values of “Assumed Familiarity”


• Ariel (1988, 1990): referring expressions indicate how accessible this piece of information is at the current stage of the discourse – “Accessibility theory”

Motivations for using Centering Theory

- CT is a processing model that relates the local utterance-by-utterance context and discourse anaphoric reference

- It is a basis to theorize about local coherence, salience and choice of referring expressions
CT: Centers & Transitions

• Centers are linguistic constructs, referents, or semantic entities that are part of the discourse model

• Each utterance has:
  – a Backward Looking Center (Cb): the most salient referent of the previous utterance that appears in the current utterance
  – a Forward Looking Center (Cf) list: a list of referents that will be projected to the following utterance
  – a Preferred Center (Cp): the most salient referent in the current utterance.

• Transitions

<table>
<thead>
<tr>
<th></th>
<th>Cb (Un) = Cb (Un-1)</th>
<th>Cb (Un) ≠ Cb (Un-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cb (Un) = Cp (Un)</td>
<td>Continue</td>
<td>Smooth-Shift</td>
</tr>
<tr>
<td>Cb (Un) ≠ Cp (Un)</td>
<td>Retain</td>
<td>Rough-Shift</td>
</tr>
</tbody>
</table>
a. Terry really goofs sometimes.
b. Yesterday was a beautiful day and he was excited about trying out his new sailboat.
c. He wanted Tony to join him on a sailing expedition.
d. He called him at 6 am.
e. He was sick and furious at being woken up so early.
Applications of CT

• Algorithms
  – Brennan et al. 1987
  – Walker 1989

• Applied to different OLs
  – Italian: Di Eugenio 1998
  – Turkish: Turan 1995
  – Chinese: Qinan 2008

• Anaphora resolution
Challenges with OL frameworks

• Theoretical approaches designed to study written language

• SL has features of spoken language (face-to-face interaction)

• Utterance boundaries:
  – Prosodic and interpretive cues
  – Extension of role shift
  … and intuition
Our application

- LSC data, but useful for other SLs
- General annotation (Nonhebel et al. 2004)
- 3 linguistic tiers added
3 linguistic tiers added
Annotation

1. Coreference tier:
   Referring expressions realized in that utterance (list of Cfs)

2. Grammatical function / category tier:
   - Subject, direct object or indirect object
   - NP, CL, pronoun, null...

3. Centering transitions:
   Backward Looking Center; Forward Looking Center; Preferred Center;
   Type of transition

   - Role tier: same index number as the coreference number for a specific referent
Referring expressions in LSC

- Full NPs & inferrables
- Pronouns & index signs
- Classifier constructions
- Verb agreement
- Null arguments
- (Role shift)
Classifiers as referring expressions

• Hypothesis of categorization:
  (i) instrumental
  (ii) entity & limb
  (iii) handling

• CL + topicalized NP which can occur in the same sentence or in some previous sentences

• CL are not referential by themselves. They only keep the referent active
CL & Role Shift

• RS indicates that the point of view holder is coreferential with a referent in the previous or matrix sentence

• CL can occur in role structures duplicating the referent of the point-of-view-holder (but not compulsorily so)

• When it happens this serves to assign discourse prominence to the referent associated with both
Idiosyncracy of Role Shift

- Role shift is usually a bridge from 3rd person to 1st person and maps anaphoric elements onto pseudo-deictic elements.

- RS is a bridge from activation to saliency (to be tested empirically).
The hierarchy of SL referring expressions is quite similar to that proposed by Ariel, Gundel at al. and Prince:

- Full NPs
  - Entity & limb CL
- Pronouns / verb agreement
  - Role shift
- Null arguments

Low accessibility markers:

High accessibility markers:
Future work

• The ranking for the Cf in a topic-prominent language
• The role that simultaneous constructions play in the accessibility scale
• Different categorisation of CL and the precise place in the accessibility scale
• The role that the use of space plays in the accessibility scale
Thanks for your attention!!

gemma.barbera@upf.edu

guillem.masso@upf.edu
References


Crasborn, O. 2007. How to recognise a sentence when you see one. *Sign Language & Linguistics* 10-2: 103-111


