Annotating disfluencies in a sign language corpus

Gaurav Mathur /gaurav.mathur@gallaudet.edu
Gene Mirus / gene.mirus@gallaudet.edu
Gallaudet University

Sign Language Corpora Network: Annotation
Stockholm
June 16, 2010

Problem

Keating, Edwards and Mirus (2008)

Disfluencies

Fillers (e.g., discourse markers, hesitations)
Slips of the hand
Repairs
  - lexical level
  - sentential level

Distinct from lexical variants and results of phonological processes

Why annotate disfluencies?

a) Capture natural discourse as much as possible
  - slips of hand in ASL (Klima and Bellugi 1979)
  - conversational repairs in ASL (Dively 2000)
  - slips of hand and repairs in DGS (Hohenberger, Happ and Leuninger 2002)

b) Mark disfluencies clearly so they can be removed ('cleaned up') if needed for a research question

c) Use for research purposes
  - frequency (identification of genre, style of presenter)
  - psycholinguistics (language production and L2 acquisition)
  - cross-linguistic comparisons
  - cross-modal comparisons

Principles for developing guidelines

Be in line with fields that have looked at disfluencies, e.g.
  - discourse analysis
  - conversation analysis
  - sociolinguistics
  - psycholinguistics
  - linguistic anthropology

Be in line with best practices for secondary processing (Johnston 2010 and Schembri and Crasborn 2010)
Annotating fillers

Guideline: establish ID-gloss for filler

Examples: UMM
WHAT'S-THE-WORD

Rationale: form and meaning conventionalized across tokens

Comment: More conventionalized than gestures

Annotating slips of hand

Guideline: use ID-gloss of intended sign and mark with tag (D for disfluency) in special tier for marking disfluencies

<table>
<thead>
<tr>
<th>ID-Gloss</th>
<th>SICK</th>
<th>BORED</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Disfluency</th>
<th>D</th>
<th>D</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Phon. Trans.</th>
<th></th>
<th></th>
</tr>
</thead>
</table>

Annotating slips of hand

Comments:

a) Cf. G:how-stupid-of-me (Johnston 2010) Indicate original intended meaning, but through ID-gloss

b) Cf. FS:WORD(WOR) (Johnston 2010) Show actual form, but through video and phonetic transcription tier


TWO MAN BOY SPAT
‘There were two men - boys - in a spat.’

Annotating lexical repairs

Guideline: use ID-gloss of the sign and mark with tag (R for repair) in special tier for marking disfluencies

<table>
<thead>
<tr>
<th>ID-Gloss</th>
<th>MAN</th>
<th>BOY</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Disfluency</th>
<th>D</th>
<th>R</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Phon. Trans.</th>
<th></th>
<th></th>
</tr>
</thead>
</table>
Repair - sentential

‘My home village’s name is vi [incompletely fingerspelled English word ‘village’], no, my home village’s name is [fully fingerspelled name of the village].’

(from Dively 1998: 155-156)

Annotating sentential repairs

Guideline:
- use ID-glosses for all signs as usual
- on disfluency tier, divide utterance into parts
  - 1st part marked D (= part being repaired)
  - 2nd part marked R (= sentential repair)

<table>
<thead>
<tr>
<th>ID-gloss</th>
<th>POS.1</th>
<th>HOME</th>
<th>FV/TV</th>
<th>NEG.REMARKABLE</th>
<th>POS.2 HOME</th>
<th>FV/TVABD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disfluency</td>
<td>D</td>
<td></td>
<td></td>
<td>R</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments:

- Can use different tags to distinguish further types of repairs

Summary

Fillers
- use ID-gloss
- no prefix needed since ID-gloss already exists

Slips of the hand
- use ID-gloss of intended sign
- mark with tag (S) in disfluency tier

Repairs
- use ID-gloss
- mark with tag (R) in disfluency tier

References


