

Semi-Automatic Identification of Semantic Relations in a Swiss German Sign Language Lexicon

Sarah Ebling, Katja Tissi, Martin Volk

{ebling|volk}@cl.uzh.ch, katja.tissi@hfh.ch

LREC Sign Language Workshop | May 27, 2012

Procedure: Two Steps

- semi-automatically obtained semantic relations in Swiss German Sign Language (Deutschschweizerische Gebärdensprache, DSGS)
- used keywords including the gloss to represent each sign
- identified semantic relations in a two-step process:
 1. automatically generated a set of pairs of semantically related entries
 2. filtered set through manual screening by a native signer

Swiss German Sign Language

- sign language of the German-speaking area in Switzerland
- approx. 6000 users
- five dialects
- our focus: Zurich dialect
- mouthings from Standard German

Step 1: Automatic Identification

- used glosses and keywords available for each sign in a DSGS lexicon (Boyes Braem, 2001) as indicators of semantic concepts
- searched for these words in GermaNet, a lexicographic reference database for German word senses containing annotations of semantic relations^a
- sample output of our approach, six antonyms of the sign ALT ('old'):

1. JUNGTER, JUNGES, JUNG ('young animal')
 $\cdot \square_{14} \emptyset \square_{\bullet}^{(\uparrow \cup \rightarrow \downarrow \downarrow)} \cdot \hat{\square}_{17} \emptyset^{(\cup)}$
2. FRISCH, NEU ('fresh', 'new')
 $\hat{\square}_{15} \emptyset \square_{\bullet}^{(\uparrow \rightarrow \downarrow \downarrow)} \square_{17}^{[1]}$
3. JUNG, JUGENDLICHER, JUGEND ('youth')
 $\bar{\square} \setminus \square_{\bullet} \prec \emptyset \square_{\bullet}^{(\chi \square \sim)} \zeta +$
4. KLEIN, JUNG, KLEINES, JUNGES ('small', 'young')
 $\cdot \hat{\square}_{17} \emptyset \square_{\bullet}^{\uparrow} +$
5. NEU, BRANDNEU ('new', 'brand new')
 $\hat{\square}_{15} \emptyset \square_{\bullet}^{(\uparrow \rightarrow \downarrow \downarrow)} \square_{17}^{[1]}$
6. NEU ('new')
 $[\square \setminus \bar{\square}_{15} \prec \square \setminus \bar{\square}_{17}] (\cup \square) [\uparrow \zeta (\downarrow \sim \chi) \rightarrow \downarrow \downarrow]$

- identified 6980 relations (cf. table to the right)

^a<http://www.sfs.uni-tuebingen.de/lsd/>

| Sub-type | Abs. | Percent. |
|----------------------|------|----------|
| Hyponym/ Hypernym | 5435 | 77.87 |
| Synonym | 817 | 11.70 |
| Meronym | 279 | 4.00 |
| Antonym | 277 | 3.97 |
| Pertainym | 59 | 0.85 |
| Participle | 40 | 0.57 |
| Related-to | 39 | 0.56 |
| Causation | 34 | 0.48 |

Step 2: Manual Screening and Results

- manually filtered the semantic relations retrieved automatically during Step 1
- task carried out by a native signer (member of our project)
- presented her with 500 randomly selected statements of the kind displayed below and asked her to rate them as True or False
 - LAND ('country' $\text{Land} \rightarrow \text{ALGERIEN}$) is a hypernym of ALGERIEN, ALGERISCH ('Algeria, Algerian' $\text{ALGERIEN} \rightarrow \text{ALGERISCH}$). ✓
 - GRUND, UMGEBUNG, MOTIV ('cause, reason' $\text{GRUND} \rightarrow \text{UMGEBUNG}$) und LANDSCHAFT, UMGEBUNG, GEGEND ('landscape, neighbourhood' $\text{LANDSCHAFT} \rightarrow \text{UMGEBUNG}$) have the same meaning. ✗
- result: 302 out of 500 statements rated as True (60.4%), 198 statements rated as False (39.6%)
- statements rated as False: relations that do not apply for DSGS → two possible reasons:
 - relation is valid for German but not for DSGS → difference in the semantic concepts of German and DSGS; examples:
 - German: *Künstler, Künstlerin, Kunstschafter, Kunstschafterin* ('artist') is a hypernym of *Musiker, Musikerin, Musikant, Musikantin* ('musician') ↔ DSGS: words are restricted to the meaning of a visual artist (i.e., a painter)
 - DSGS: TRAINING, TRAINIEREN ('practice') is confined to the domain of physical activity ↔ German: may involve any sort of training
 - relation is also not valid for German; e.g., false statement due to an ambiguity of the word *Vater* (which can mean both 'Holy Father' as well as 'father')

Identifying Sub-Types

- automatically extracted sub-types of the same type within the concept of double glossing (Konrad, 2011)
- looked at pairs of hyponyms of the same hypernym that are form-equivalent (have the same HamNoSys notation)
e.g., BACH ('brook') and FLUSS ('river') have same form ($\text{Q}_{\Delta 0}^{\text{I} \sim}$) and same GermaNet hypernym, *Gewässer* ('stretch of water')



- heuristic: two form-equivalent signs that are semantically related have the same underlying image and image producing technique
- identified 53 sub-type pairs of this kind

References

Penny Boyes Braem. A multimedia bilingual database for the lexicon of Swiss German Sign Language. *Sign language & Linguistics*, 4:133–143, 2001.

Reiner Konrad. 2011. *Die lexikalische Struktur der Deutschen Gebärdensprache im Spiegel empirischer Fachgebärdenlexikographie. Zur Integration der Ikonizität in ein korpusbasiertes Lexikonmodell*. Gunter Narr Verlag.

