Towards Fair Licenses for Data from the DGS Corpus Project

Thomas Hanke, Sung-Eun Hong, Susanne König, Gabriele Langer, Rie Nishio, Christian Rathmann
University of Hamburg, Institute of German Sign Language and Communication of the Deaf

GERNERAL INFORMATION
Data Collection
• 2009-2011
• 12 locations (mobile studio)
• 328 informants balanced for sex, age, region
• Studio session:
  • 2 informants (peer-to-peer situation)
  • Moderated by Deaf contact person
• Duration: one day (approximately 5.5 hours of filming plus breaks)
• Elicitation tasks (about 20 different tasks)
The Public Corpus
• Selected parts of the reference corpus (approx. 50 hrs) will be made publicly accessible (including English translation and basic transcription/annotation)

Data Sharing Options
• Informed Consent
• Selected part of the corpus (part of the signed movies and their transcriptions) will be shared with other researchers (interested in more data by informant)
• Contact data may be shared with other researchers (interested in more data by informant)

Informed Consent
• Purpose of research
• Data collection: connect, talk, video
• Use of data within the project
• Data sharing options
• Informed Consent
• Signed Consent

DGS Consent Procedure
• Use in our own research project
• Sharing & Options

License Options Chosen (122 out of 320)

Approval and Exclusions
• Total number of exclusions: 60
• Total time of excluded material: 0.48 hrs of estimated 125 hrs material (0.64 %)

Where do exclusions occur?
• In-between tasks, during breaks (9)
• within “free conversation task” (23)
• within “Deaf events task” (8)
• other tasks (20)

Reasons for exclusion:
• Content-related:
  • Names, name signs (22)
  • Opinions on other persons, groups, topics
  • Information about other persons, institutions
  • Information about own person, workplace
  • Private communication in between tasks
• other
• Related to studio situation (e.g. questions on task, feedback comments)
• Non-linguistic reasons (e.g. blowing nose, fixing clothing)

Poster presented at the Workshop on Exploitation organized by the Sign Linguistics Corpora Network, held in Berlin, Germany Dec 3-4, 2011.