Designing Elicitation Stimuli and Tasks for the DGS Corpus Project

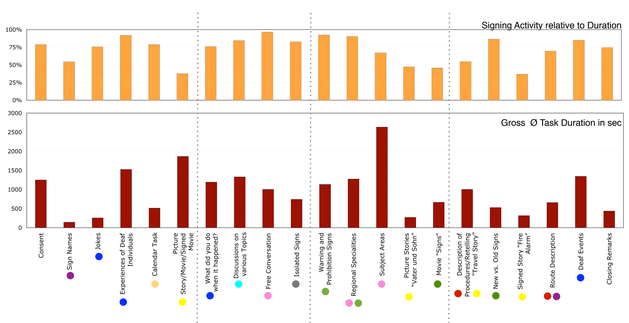
Thomas Hanke, Sung-Eun Hong, Susanne König, Gabriele Langer, Rie Nishio, Christian Rathmann
University of Hamburg, Institute of German Sign Language and Communication of the Deaf

Background Information

The DGS Corpus Project is a long-term project with two major aims: (i) to establish an extensive corpus of DGS and (ii) to develop a comprehensive dictionary of DGS-German based on the analysis of the corpus data.
The purpose of the corpus is to document the use of DGS and to provide material of and on Deaf culture and life. The corpus will be a resource that can be used for a variety of research questions, which is why it needs to consist of a large variety of discourse modes and grammatical structures as well as various subject areas. As one of the project aims is to compile a general dictionary of DGS, the corpus should also provide enough material on the lexicon of DGS and its use.
To this aim we have to make sure that different text types are collected. Our data collection (Nishio et al. 2010) consists mainly of staged communicative events. Consequently, stimuli had to be developed in order to evoke different types of communicative events, such as monologues, dialogues, emotional vs. factual text, re-tellings, prepared vs. fully spontaneous text etc.
To achieve this goal, stimuli have been developed, tested, modified, tested again – and sometimes been dismissed.
From Jan. 2010 to Dec. 2011 the data is collected from more than 300 informants in a mobile studio which is set up at 12 sites throughout Germany. The tasks are presented to the informants on screen. In order to keep it simple for the regional contact person, who also moderates the data collection, and to make sure all informants get the same instructions, for each task movies with explanations in DGS are provided.

Formats	Tasks (examples)	Goal
● narration	Jokes (prepared signing)	Deaf cultural heritage
● re-telling	Frog Story (pre-structured signing by fixed order of pictures)	Cross-linguistic research
● argumentation	Discussion on various topics	Emotional output
● isolated signs	Isolated Signs	Regional variations
● description	Description of Procedures	Description of different activities
● explanation	Sign Names	Deaf cultural heritage
● negotiation	Calendar Task	Signs for numbers and days of the week
● free conversation	Free Conversation	Conversation without moderator
● discussion	Warning and Prohibitive Signs	Expressions of negation

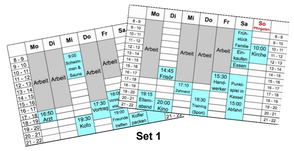
Task duration
The combination of different formats also allows us to mix tasks that require some time for introducing what the signers are supposed to do with tasks where almost no instruction time is needed. On the basis of the logs from 46 data collection sessions completed so far, the signers are actively communicating with each other during 4.14 out of 5.6 hours (74.0%). The ratio ranges from 37.0% (retelling) to 96.8% (free conversation).

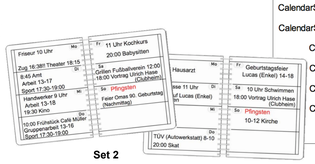


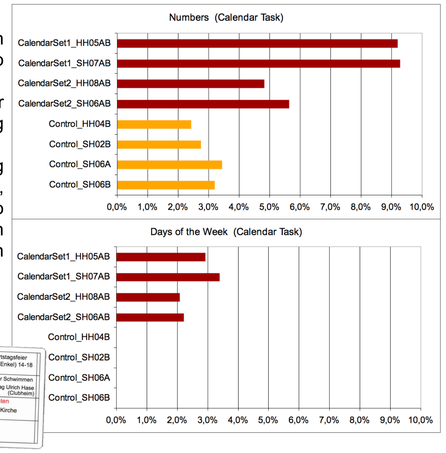
Effectiveness of Tasks – A Preliminary Study

In order to make sure that less-frequent grammatical constructions and vocabulary are covered in the corpus in large enough quantities, some tasks were defined with the goal to augment the frequency of these target constructions while staying within the paradigm of corpus collection. In order to measure the effectiveness of these tasks, we compare the relative frequency of target constructions from theses tasks with that from a close-to-spontaneous narrative signing task ("What did you do when it happened?"). At this point of time where corpus transcription has only started, the frequency is calculated as the absolute number of target hits divided by the total signs count in the task as extrapolated from per-signer signing speed sampling.

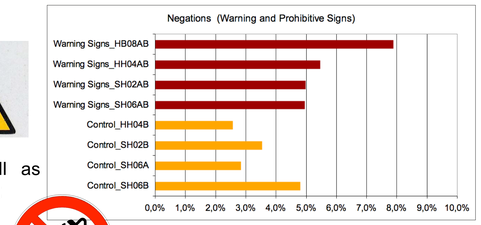
Calendar Task
Task: Informants are shown a one-week calendar with fictive appointments and are instructed to arrange two meetings of two hours respectively.
Goal: Collecting signs for days of the week, number signs (e.g. for time terms), activity terms; eliciting planning and negotiation discourse.
Results: The task succeeded in effectively eliciting signs for *numbers* and *days of the week*. For the latter, having tasks like this even might be necessary as no days of the week appeared in the control sample. In addition, calendar layouts seem to have an effect on how the conversation focuses on time and days.



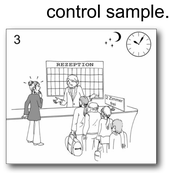


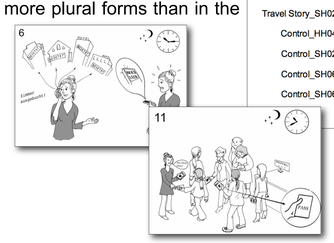


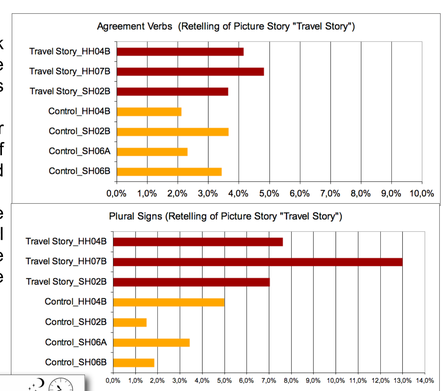
Warning and Prohibitive Signs
Task: Informants look at 16 warning and prohibitive signs collected from different places of the world (and therefore unfamiliar to them) and discuss what they might possibly mean.
Goal: Eliciting negative expressions (as well as expressions of possibilty, judgement and opinion); warming up by having fun.
Results: The task elicited *negations* effectively, although the difference to the control group is not as large as we had hoped for.



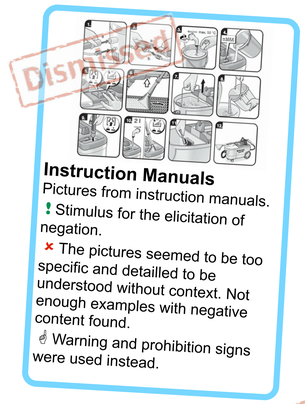
Travel Story
Task: The travel story is a picture story re-telling task on a bus journey chaperoned by a travel guide. One informant looks at the story consisting of 17 scenes and re-tells it to the second informant in 7 sections.
Goal: Eliciting various ways of spatial use for directionality and plurality, eliciting many instances of already known directional verbs such as GIVE and ASK in context.
Results: *Agreement verbs* appeared slightly more often in re-tellings of the travel story than in the control group. As for *plurality*, the stimulus proved to be more successful, i.e. there are more plural forms than in the control sample.

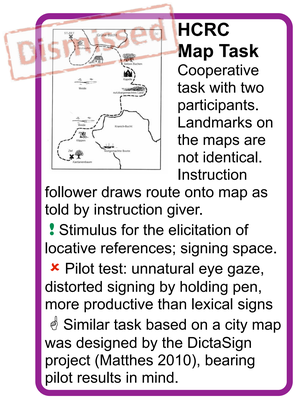


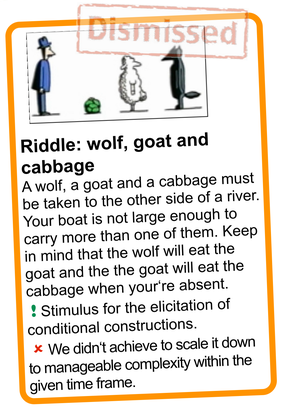




Abandoned Tasks







Criteria for Future Stimuli Design

One criterion for our task evaluation was the ratio of invested explanation and stimulus-viewing time compared to sign production time. To ensure that there will be enough signed output, very complex tasks that needed a long explanation and stimulus viewing time, or that were found to produce many questions on the task by the informants had to be left out or simplified. This was also necessary in order to prevent too much signing to and by the moderator, who is supposed to stay passive as much as possible. We also wanted to keep it simple for the moderator and to ensure that the informants stay at ease at all times and do not feel overstrained or being tested.
•All of the tasks were tested with deaf subjects and some of them had to undergo several rounds of modifications while others did not work out at all and thus were not included for the actual data collection.
From our findings there are some points to keep in mind for future stimuli design:
• Be aware of cultural differences (iconography, school curricula, customs/traditions…).
• Be aware that the required register for a given task might be non-existent in the target language – or might be highly unusual for informants to use.
• Be aware that meta-linguistic knowledge cannot be taken for granted.
• Be cautious about giving tasks which could evoke a testing situation.

References

Anderson, H. et al. (1991): The HCRC Map Task Corpus. In: Language and Speech 34, 351-366.

Blanck, Dolly / Hanke, Thomas / Hofmann, Ilona / Hong, Sung-Eun / Jeziorski, Olga / Kleyboldt, Thimo / König, Lutz / König, Susanne / Konrad, Reiner / Langer, Gabriele / Nishio, Rie / Rathmann, Christian / Vorwerk, Stephanie / Wagner, Sven (2010): The DGS Corpus Project. Development of a Corpus Based Electronic Dictionary German Sign Language – German. Poster presented at the Theoretical Issues in Sign Language Research (TISLR) 10 Conference, Sept 30 - Oct 2, 2010 at Purdue University, Indiana, USA.

Matthes, Silke / Hanke, Thomas / Storz, Jakob / Efthimiou, Eleni / Dimiou, Nassia / Panagiotis, Karioris / Braffort, Annelies / Choisier, Annick / Pelhate, Julia / Safar, Eva (2010): Elicitation Tasks and Materials designed for Dicta-Sign's Multi-lingual Corpus. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, LREC 2010, 22-23 May 2010, Malta. pp. 158-163.

Nishio, Rie / Hong, Sung-Eun / König, Susanne / Konrad, Reiner / Langer, Gabriele / Hanke, Thomas / Rathmann, Christian (2010): Elicitation methods in the DGS (German Sign Language) Corpus Project. In: Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, LREC 2010, 22-23 May 2010, Malta. pp. 178-185.

Poster presented at the Theoretical Issues in Sign Language Research (TISLR) 10 Conference, Sept 30 - Oct 2, 2010 at Purdue University, Indiana, USA.
The research leading to these results has received funding from the German Academies of Science Programme.