



The current status of GFN (German FrameNet)

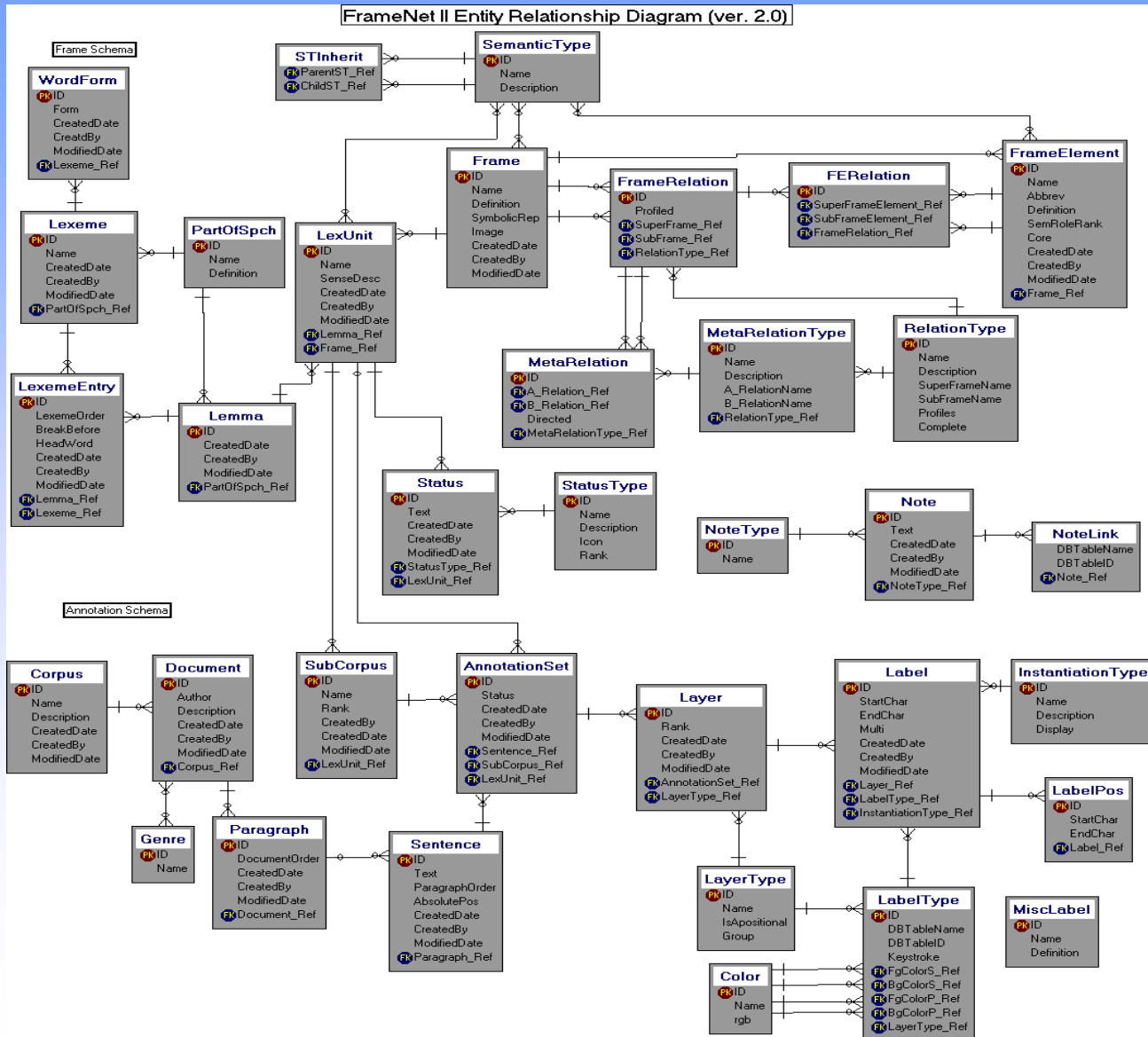
Hans C. Boas, Elias Ponvert, Mario Guajardo, and Sumeet Rao
University of Texas at Austin
<http://gframenet.gmc.utexas.edu>

Goals of Project

1. Take English FrameNet database
2. Cut out all English-specific information, leaving in place conceptual information (frames, FEs, relational information, etc.)
3. Follow Berkeley's procedures for repopulating the database with German-specific information (using the FN desktop)



1. Take English FN database



2. Cut out all English-specific info

- Downloaded Berkeley's FN database installed multiple copies on server for testing purposes.
- Cut out northwest area of the database: Lexical Unit, POS, Lexeme, WordForm, etc.
- Cut out southwest area of the database: Corpus, Document, Paragraph, etc.



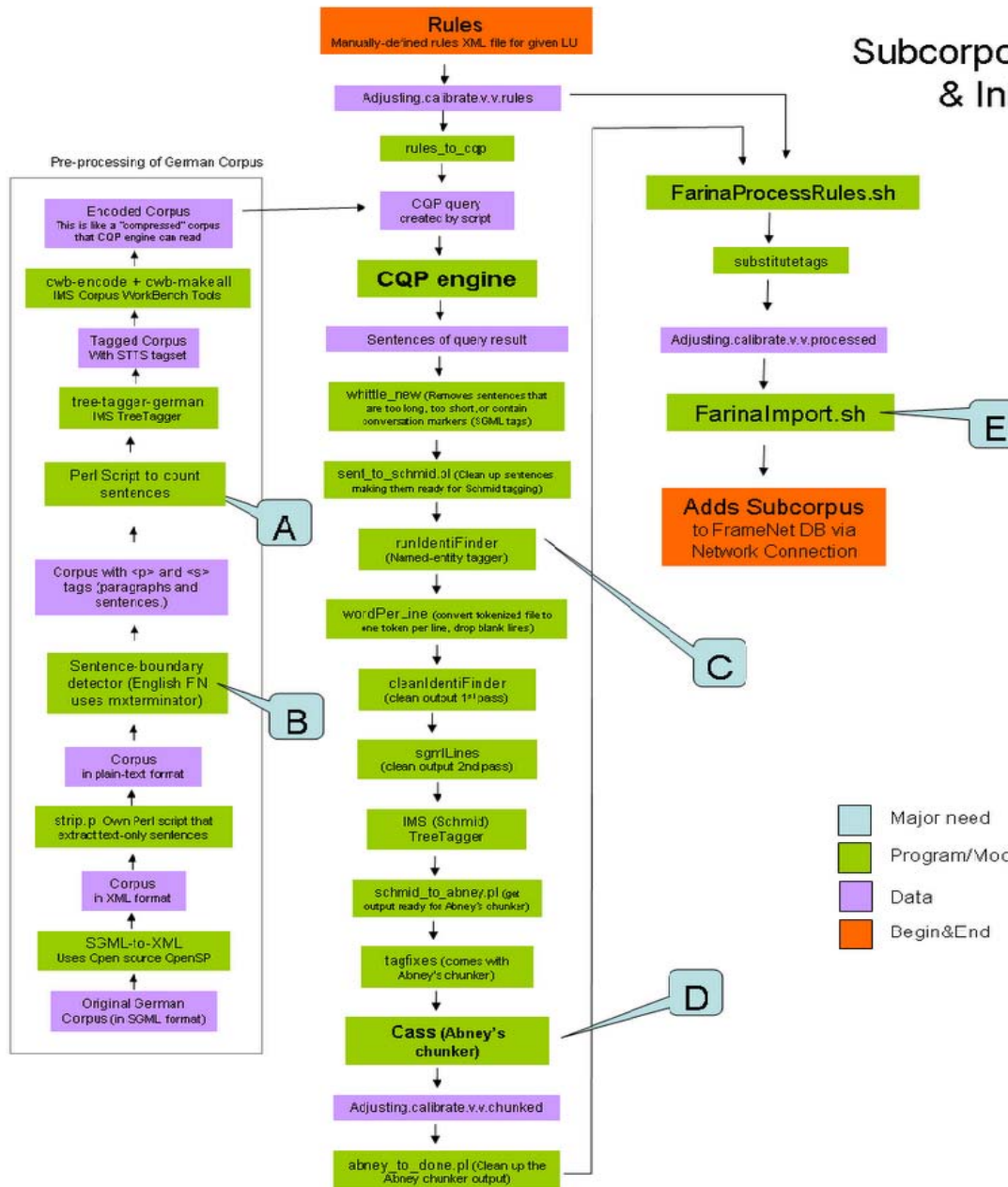
3. Re-populating the database with German-specific information

- German corpora for GFN:
 - LDC German newspaper corpus
 - Institut für Deutsche Sprache: Datenbank gesprochenes Deutsch
 - Tagging with STTS



Subcorporation Pipeline & Information Flow

May 22, 2006



How do we get there?

The to-do list ...

- German Chunker
- Sentence Boundary Disambiguator (Satz, Uplug)
- Named Entity Recognizer
- Write German “rules”
- Streamlining the “pipeline” for GFN



Some thoughts on ...

- Funding (UT, NSF)
- Working with students from different departments
- Multilingual NLP efforts at UT
- Collaboration with SFN, JFN
- Interlingua?



Developing FNs for more languages?

- Develop infrastructure
- Collaboration between teams
- Providing information, in particular list of known problems and how to solve them
- GFN blog:
<http://gframenet.gmc.utexas.edu>



Process_End

- *come to a close.v, come to a end.v, conclude.v, draw to a close.v, draw to a end.v, end.v, end_cessation.n, end_final-subevent.n, final.a, finale.n, pass.v*

