



www.typecraft.org

It is said that a language is a dialect with an army and a navy; in today's Information Society a language is a dialect with a dictionary, a grammar, a parser and a multi-million word corpus of texts. When you've got all of that, get a speech database, and your language is set to compete in the age of Globalization.¹

In recent years the number of digital language tools has increased rapidly. Huge databanks of linguistically annotated natural language text have been built, for example for American and British English. Yet, so far many of these initiatives remain restricted to a few key languages,



since the usage and the maintenance of digital tools for language documentation often requires expertise and access to resources which communities with minority or endangered languages lack. To avoid that (national) language resources are out-sourced, language documentation tools must become a more general commodity. With 'TypeCraft' we present an Glossing Editor combined with a natural language database and a wiki – which serves as a communication and knowledge sharing tool.

With TypeCraft the user glosses (tags or annotates) text on the word and sentence level. The main purpose of annotation is to illuminate meaning and grammatical structure, and thus to enhance and document our understanding of language.

TypeCraft makes annotating 'easy', 'distributive', 'interactive' and 'software independent'. With TypeCraft online linguistically annotated data becomes generally available. Here are some of the salient features of TypeCraft:

- ◆ import of text and automatic sentence break-up
- ◆ easy-to-use tabular interface for sentence glossing
- ◆ Lazy Annotation Mode (sentence parsing)
- ◆ drop-down-style reference lists, composed of widely used linguistic symbols for word and construction level annotation
- ◆ internal knowledge sharing facility, allowing information exchange between individuals and projects
- ◆ privacy combined with data sharing facilities implemented as several different work spaces
- ◆ export of annotated natural language data to Microsoft Word and Latex as well as export of XML for further processing

¹ Free after Ostler 1998