

WikiTrans: Swedish-Danish Machine Translation in a Constraint Grammar Framework (invited talk)

Eckhard Bick

Institute of Language and Communication
University of Southern Denmark
eckhard.bick@mail.dk

1 Abstract

This talk presents an MT system for the automatic generation of Danish Wikipedia articles from Swedish originals. The translated Wikipedia (WikiTrans) is indexed for both title and content, and integrated with original Danish articles where they exist. Newly added or modified articles in the Swedish Wikipedia are monitored and handled on a daily basis. The translation approach (GramTrans) uses a grammar-based machine translation system with a deep, structural source-language analysis. Morphosyntactic disambiguation and lexical transfer rules exploit Constraint Grammar tags and dependency links to access contextual information, such as syntactic argument function, semantic type and quantifiers. Out-of-vocabulary words are handled by derivational and compound analysis with a combined coverage of 99.3%, as well as systematic morpho-phonemic transliterations for the remaining cases. Reflecting the similarities between Swedish and Danish, the system achieved high BLEU scores (0.65-0.8 depending on references), and outperformed standard STMT and RBMT competitors by a large margin.

2 Biography

Dr. Eckhard Bick is a computational linguist and project leader for the VISL lab at the University of Southern Denmark, where he works as a language technology researcher at the Department of Language and Communication (ISK). Over the years he has designed and developed grammars, corpora, lexical resources and applicational tools for a large number of languages, including most of the Romance and Germanic languages. Eckhard Bick is a leading expert in the field of Constraint Grammar, with a current focus on semantic annotation and machine translation. Eckhard Bick has published extensively on various aspects of computational linguistics and participated in a large number of international research projects.