

[From: CLAW '96, Universiteit Leuven]

A Path from Sublanguage to Controlled Language

Richard Kittredge
University of Montreal *
and
CoGenTex, Inc.

Abstract

The notion of *natural sublanguage*, emphasizing a spontaneous language subsystem having tight restrictions on usage of domain-dependent word classes, contrasts with the separate phenomenon of controlled language, where the conscious introduction of a broad set of writing guidelines may span a whole text genre. By its focus on a single standard to serve a disparate community of users, CL engineering may give results that are not optimal for some classes of users, or for particular sublanguages. The proposal here is to use better sublanguage models to make the design of CLs more rational, and eventually to allow a more flexible and tailored set of standards. Current work in natural language generation (NLG) already faces issues of whether and how to "normalize" output. Work in multi-lingual text generation is finding subtle differences in rhetorical structures and other features across languages that should affect flexible CL design. As NLG of technical manuals becomes more practical, these concerns will strongly intersect those of the CL community. To build better CLs, it is important to collect and analyze cases where CL features (e.g., lexical substitutions) seem confusing to some CL users or alien to SL experts. A better modelling of the reference SLs and of the language proficiency of user groups should be supported by workstation tools for CL, and will pave the way for a controlled introduction of NLG techniques in service of CL concerns.

*Department of Linguistics and Translation, CP 6128, Montreal, QC H3C 3J7, Canada.
E-mail KITTREDG@IRO.UMONTREAL.CA.