

Reports from the SCARRIE project
Editor: Anna Sågvald Hein

Two Reports on CORRIE for SCARRIE

Papers by:

Erik Tjong Kim Sang
Leif-Jöran Olsson



Reports from the SCARRIE project
Editor: Anna Sångvall Hein

Two Reports on CORRIE for SCARRIE

Papers by:

Erik Tjong Kim Sang
Leif-Jöran Olsson

WP CL&LE 3

Preface by the editor

SCARRIE is short for *Scandinavian Proof-reading Tools*. It is the name of a project within the EU TELEMATICS APPLICATIONS Programme. The project run for 30 months starting in November 1996.

"The SCARRIE project aims at the development of a high-quality proof-reading tool for the Scandinavian publishing industry. A concrete concrete result of the project will be a carefully evaluated demonstrator, the SCARRIE pilot, designed to meet the needs formulated by a user group consisting of representatives for Danish, Norwegian and Swedish newspapers and publishing houses. " (from LE3-4239 SCARRIE Project Programme. Annex I.)

The SCARRIE demonstrator for Swedish was developed by the Department of Linguistics at Uppsala university. Fundamental user input was provided by two Swedish newspapers, Svenska Dagbladet and Upsala Nya Tidning. We gratefully acknowledge their contributions. See further <http://www.scarrie.com/> for general information about the project and <http://stp.ling.uu.se/~ljo/scarrie-pub/> for a test version of the resulting Swedish SCARRIE pilot.

The achievements made in the SCARRIE project were continuously delivered to the European Commission via the project officer. When the project was still running the availability of most of the deliverables was restricted. Now the project is closed, and we make some of them available to a larger community. They all concern the development of the Swedish SCARRIE pilot.

The SCARRIE word checker is based on an extended version of a Dutch software CORRIE. Both papers presented below relate to CORRIE. In the first one the results of the initial testing of CORRIE with respect to a number of features are presented. The second one concerns error correction. Error correction in CORRIE is based on both graphemic and phonemic representations of potential errors. The second paper deals with the conversion of text strings from a graphemic to a phonemic representation. In particular, the formulation of conversion rules in the CORRIE formalism is focused.

For reports on other aspects of the SCARRIE project, see Working Papers in Computational Linguistics & Language Engineering No. 4 - No. 13.

Contents

<i>Testing CORRIE for SCARRIE, Deliverable 1.2.</i> Uppsala, Nov 1997. Tjong Kim Sang, Erik.....	1
<i>Specification of Phonemic Representation, Swedish, Deliverable 4.1.3.</i> Uppsala, Jan 1998. Olsson, Leif-Jöran.....	23