Writing for Language-Impaired Readers

Aurélien Max

LIMSI-CNRS & Université Paris Sud
Bâtiment 508, F-91405 Orsay Cedex, France
`aurelien.max@limsi.fr`

Abstract. This paper advocates an approach whereby the needs of language-impaired readers are taken into account at the stage of text authoring by means of NLP integration. In our proposed system architecture, a simplification module produces candidate simplified rephrasings that the author of the text can accept or edit. This article describes the syntactic simplification module which has been partly implemented. We believe the proposed approach constitutes a framework for the more general task of authoring NLP-enriched documents obtained through validations from the author.

1 Introduction

The goal of NLP, as it seems, is mainly to do some processing on existing text. Another domain of application that we believe has a great potential is the use of NLP during text creation, where it can help authors write better documents in a more efficient way. A lot of the difficulties when processing real-life documents arise from the inherent complexity of natural language, which requires word-sense and syntactic structure disambiguation, to name just two. In fact, rule-based and statistical NLP systems are rather good at finding hypotheses, but they often fail when it comes to ranking them and finding the appropriate solution in context.

Some cases can certainly justify the extra cost of annotating the text with the result of the correct analysis, thus permitting much better results on NLP tasks. This concept has already been investigated, for example in the Dialogue-based Machine Translation paradigm [1] whereby a monolingual writer answers ambiguity questions. This process yields a disambiguated analysis for each sentence that is then sent to a machine translation engine.

The kinds of annotation that can be obtained through interaction can be of very different natures. One kind is a transformation of the initial text: for example, annotations at the paragraph level can be assembled to constitute a summary. Transformations at the sentence level include paraphrasing and its differents uses. Among them, text simplification has attracted significant interest in the past years [4, 3, 5]. The most notable application of text simplification has been as an assistive technology for people suffering from aphasia, a loss of language that can result in severe comprehension disorders. It is very unlikely that a text transformation system could produce a coherent text conveying the