

Constructing readable language: How well can a computer create documents for people with poor reading skills?

Sandra Williams, Department of Computing Science, University of Aberdeen,
October 2005

Poor literacy is a big problem in the U.K. where one in five adults has problems with reading. It would help these people if documents were easier to read. Natural Language Generation (NLG) is the field of Artificial Intelligence concerned with computer generation of documents from underlying data. NLG systems make choices about document content (what to say) and linguistic expression (how to say it). This seminar describes research into what is required to enable an NLG system to make choices that improve readability, work that crosses boundaries between AI, linguistics and psychology. I describe discourse-level choices for readability that were the subject of my PhD and the slightly broader range of choices, including lexical choice, that are being investigated in our current project, SkillSum.

I will describe the NLG system that was built for this research in collaboration with basic skills experts. The NLG system forms part of a web-based basic skills screening application that adults can use to check their literacy and numeracy skills. It generates personalised feedback for people who use the screener. It uses their screener results to construct a report that describes their strengths and weaknesses and gives advice on how to get help, if necessary.

The NLG system was built using empirical linguistic knowledge acquired from two corpora: a corpus annotated with discourse relations and a corpus used to obtain written and spoken lexical and phrase frequency data. The system makes decisions about linguistic expression based on rules derived from psycholinguistic data.

We evaluated the NLG system to find out if the linguistic choices it makes result in more readable basic skills reports. We constructed two versions of its linguistic choice rules: one based on the most common features found in written English and one based on psycholinguistic data and data from the spoken language corpus. We then conducted experiments with poor readers using the basic skills screener to test their skills and the NLG system to generate a report for them using (randomly) one of the two sets of rules. We investigated which version they found most readable and which they preferred using a number of techniques. I will describe our experiments and our results to date.