New computer-based tool measures readability for different readers

May 14 2013

Today most public services involve electronic communication, which requires that people are able to read relatively well. However, a significant number of adults cannot fully understand the texts they read for example on the Internet. A new doctoral thesis from the University of Gothenburg shows that a new model called SVIT can be used as a tool to measure the readability of texts and therefore how appropriate they are for different target groups.

'Modern life requires people to be good at navigating in society. This includes having to make many important decisions, for example when dealing with the bank or some state agency on the internet. In many of these situations, a person's success depends on his or her ability to read,' says the author of the thesis Katarina Heimann Mühlenbock.

To measure the readability of a text, Sweden has relied on a readability index called LIX since 1968. The LIX measure indicates what a text looks like on the surface, or more exactly the average sentence length and the proportion of the words that are longer than six letters.

Today more modern language technology and digital language resources are allowing for more precise readability analyses. Heimann Mühlenbock assessed previous studies on what it is in a text that affects different readers' ability to understand it. It may for example be a matter of word variation and sentence structure. The thesis investigates at which linguistic levels these features of a text are found, and whether they really differ to any measurable extent.

Heimann Mühlenbock analysed a total of 23 different linguistic variables and then used them in a computer programme designed to classify texts based on not only complexity (ordinary vs. easy-to-read texts), but also genre (fiction, news and information).

It turned out that the computer programme yielded an accuracy of 84-100 per cent; the corresponding figure for LIX was 45-99.5 per cent. The new model, which is called SVIT, outperformed LIX across the board, with a difference between the two methods of 0.2-55.4 percentage units.

'My results show that SVIT can be used both to customise texts to specific target groups and to find texts that suit a certain reader profile. This is very useful since individuals with different linguistic backgrounds need different types of texts,' says Heimann Mühlenbock.

Provided by University of Gothenburg