Simplify to understand

Tool developed by UFRGS' researchers makes texts from the health field easier to understand

First published: January 23, 2020

By: Thauane Silva

In a survey conducted last year by the NGOs Instituto Paulo Montenegro in partnership with the NGO Ação Educativa and the IBOPE Inteligência (educational division of an organization specialized in statistical analysis), it was pointed out that about 3 out of 10 Brazilians are functionally illiterate. This means that, although they can read and write, they have difficulty understanding simple texts and performing basic mathematical operations. The survey evaluated the literacy levels of the Brazilian population between the age range of 15 and 64. In addition, a survey conducted by Google showed that 26% of Brazilians use the search engine to look up their health problems before seeing a doctor.

Access the tool for free on the research group's website. Photo: Gustavo Diehl / UFRGS

Bearing this in mind, professor Maria José Finatto and doctoral student Liana Paraguassu, both from UFRGS' Institute of Languages and Literature, decided to create a digital tool called MedSimples to simplify texts related to the health field. The project started in 2016, after professor Finatto noticed the complexity of the texts from the health field she used in her translation class. Subsequently, as soon as she started her doctorate on health literacy, Liana Paraguassu also joined the Textual and Terminological Accessibility group, which includes students of Computer Science as its research members.

The MedSimples tool is made of four different databases: *CorPop*, a compilation of words from vernacular written Portuguese; *Parser Passport*, a free resource developed by Leonardo Zilio, Rodrigo Wilkens and contributors to identify the grammatical class of words; *TeP 2.0*, a digital dictionary of synonyms and antonyms to aid in the understanding of difficult words; and a glossary created by the researchers with terms from the health field, along with their definitions simplified according to the reader's profile. These databases work together to recognize terms and words that need to be made more accessible for the selected reader profile.

The tool is similar to Google Translate: the users copy part of a text found on the internet or written by themselves, insert it in the search box, and then choose the subject to which the excerpt belongs and the profile of the person who will read it (the options take into account, for example, the education level of the reader). A suggestion for textual simplification will then appear on a block at the bottom of the page, with terms belonging to the health field highlighted in green and explained between parentheses. Portuguese words considered difficult will appear with their synonyms written in blue.

The reader profiles that will be available are still to be defined by the research group, based on the report of the literacy survey. Finatto explains the report will reveal what Brazilians read, how much of it they understand, how many years of education they have, and what their reading habits are.

At first, the tool was designed for use by journalists, healthcare professionals, and translators who work with the simplification of complex texts. However, during the planning process of the project for the *Latin America Research Awards* 2019 (LARA 2019), created by Google, the researchers came across the already mentioned data collected by the company, which points out that 26% of the Brazilian population seeks information about health problems directly on the internet. "So we think MedSimples can be used by anyone," explains Paraguassu. On the research group's website, there can also be found a glossary and a compilation of simplified sentences about some subjects from the health field.

The tool is already available for use, but still at an initial stage of development, collecting terms and difficult words related to Parkinson's disease. It is also only available for one reader profile, "adults with nine years of schooling". The researchers intend to expand the tool not only to reach more reader profiles, but also to reach other subjects, such as children's health--the next coming subject--Alzheimer's disease and measles. They expect MedSimples to be more functional by the end of this year.

In November 2019, the project was one of the 25 winners of LARA 2019, and was awarded a research grant for one year. "We believe that both this tool and the work we are doing in the Textual and Terminological Accessibility group have an important social role because we live in a country where people have increasingly access to the internet and to information, but not to quality education that enables them to understand all the content that is available to them," says Paraguassu. According to professor Finatto, one of the project's goals is also the creation of a massive open online course on "Easy Text", which will be soon available on the Lúmina platform.
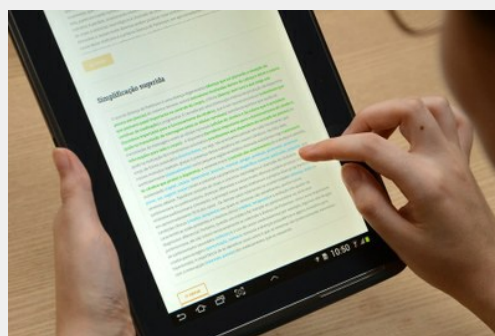
Translated into English by Caroline Winge, under the supervision and translation revision of Elizamari R. Becker (P.h.D.) – IL/UFRGS.