

## **Adapting the Sketch Engine interface for the purposes of an error corpus**

Iztok Kosem

Recently, a new feature for displaying annotated errors and corrections has been added to the Sketch Engine. Originally used by the Cambridge University Press and limited in access to the researchers working on their error corpora, the feature was part of a specially designed Sketch Engine interface and not available to regular users of the Sketch Engine. Now, after the feature was made available in the Sketch Engine, the compilers and researchers of error corpora have the option to include their corpora in this corpus tool; and considering the popularity of the Sketch Engine, as well as the fact that error corpora are still often tied to tailor-made corpus tools (which costs money and means that researchers have to learn how to use different corpus tools), this may be an opportunity for the Sketch Engine to establish itself as the tool for error corpora analysis.

This presentation will present a project in which the Sketch Engine interface for error corpora was adapted because target users, i.e. language teachers and publishers, were less proficient in using corpus tools. The error corpus that was part of the project was the 1-million-word Šolar corpus of young learner writing, built within the Communication in Slovene project and containing texts of Slovene elementary and secondary school students. The corpus contains annotated student errors, as identified by their teachers, and teacher corrections.

Considering the aforementioned target users, our aim was to design an interface for the corpus that would be user-friendly and easy to use. Cambridge University Press was used as a point of departure, with the adapted interface meeting the following requirements:

- a) display of errors and corrections (but not annotation tags) had to be the default option, with both types of information clearly distinguishable from each other, and from non-errors/corrections and the surrounding text (this was achieved by displaying errors in red, correction in green, and all other text in black);
- b) the users had to be able to reveal in the concordance output the annotation tags, which had to be named in an easily understandable manner;
- c) the tool had to be localised into Slovene with corpus terminology used only when necessary, and with tips and help provided for each function of the tool.

A short demo of the interface will be given, and changes that were needed to include the corpus in the Sketch Engine interface (e.g. error categorization) will be discussed. In addition, the presentation will show the option of further annotation of concordances, available in the Sketch Engine and not part of the Šolar corpus interface, as well as discuss some ideas on how to combine the CPA functions with error analysis and concordance annotation. Closing remarks will include suggestions on further improvements/developments of the error corpus feature in the Sketch Engine.