

Meike Meliss/Vanessa González Ribao

VERGLEICHBARE KORPORA FÜR MULTILINGUALE KONTRASTIVE STUDIEN

Herausforderungen und Desiderata

Abstract This contribution aims to show the necessity of working in the development of multilingual corpora and appropriate tools for multilingual contrastive studies. We take the corpus of the lexicographical project COMBIDIGLEX as example to show, how difficult it is to build a suitable data basis to study and compare linguistic phenomena in German, Spanish and Portuguese. Despite the availability of big reference corpora for the three languages (at least for written language), it is not able to obtain a comparable data basis from, because the mentioned corpora are created according to different requirements and they are also powered by disparate information systems and analyse tools. To break the status quo, we plead for increasing research infrastructures by means of compatible language technology and sharing data.

Keywords Corpus linguistics; comparative corpora; contrastive multilingual linguistics; language technologies

Contact information

Meike Meliss

Universidad de Santiago de Compostela

meike.meliss@usc.es

Vanessa González Ribao

Postdoc-Stipendiatin der Fritz-Thyssen-Stiftung

vanessina_gr@hotmail.com