

Invitée par l'Institut de langue et littérature allemandes

dans le cadre du cycle de conférences SciLAC

Madame Kristin Kopf

Collaboratrice scientifique
(Westfälische Wilhelms-Universität Münster)

présentera la conférence suivante :

**« Bad data + bad data = good data ?
Combining the diachronic corpora of German »**

le jeudi 1^{er} novembre 2018 à 16h00 (précises)

Lieu : Université de Neuchâtel – salle RE42
(Faculté des lettres et sciences humaines - Espace Louis-Agassiz 1)

Toute personne intéressée est cordialement invitée.

Abstract:

In recent years, many corpora of historical German have become available or been extended considerably. This now allows for quantitative analyses of phenomena across centuries by using data from multiple sources. However, combining data from different corpora has its pitfalls: what appears to be a diachronic change may be due to differences in corpus composition, tagging and/or annotation. Among the major challenges are finding comparable queries between corpora and the interpretation of diverging (and converging) results. In this talk, I will use developments that are already well-documented as a means to evaluate the corpora, presenting a series of pilot studies that combine data ranging from 1350 to 1900.