

## FALKO

### Ein fehlerannotiertes Lernkorpus des Deutschen

**FALKO** ist ein fehlerannotiertes Lernkorpus des Deutschen, welches in der aktuellsten Version auf sechs Subkorpora basiert. Diese unterscheiden sich in zwei Faktoren: Schreibaufgaben und Muttersprache. (Reznicek, Lüdeling, Krummes, Schwantuschke, 2012: 4).

Da es bereits Lernkorpora zum Erwerb der englischen Sprache gibt, soll mit dem Falko-Lernkorpus ein Korpus für den Erwerb der deutschen Sprache angelegt und folglich eine Lücke in der Fremdsprachenerwerbsforschung geschlossen werden.

Alle im Korpus enthaltenen Texte wurden unter Prüfungsbedingungen verfasst. Die sogenannten Kontrollkorpora wurden ebenfalls unter den gleichen Bedingungen und mit denselbigen Anforderungen erhoben wie die Lernkorpora (ebd.).

Zu den **Metadaten** im Korpus lässt sich folgendes feststellen:

Neben den gängigen Faktoren wie Geschlecht, Alter und Muttersprache werden auch Faktoren wie der Schulabschluss, die bereits erlernten Fremdsprachen mit der zeitlichen Spanne und möglichen Auslandsaufenthalten angegeben, sowie die Ergebnisse des C-Test (Einstufungstest) und die Korrektoren der Tests angegeben (ebd.: 9ff.).

Für den Bereich der **Annotationen** lässt sich des Weiteren folgendes konstatieren:

Für alle Wortarten wurden die Wortarten (POS) sowie die Lemmata (lemma) automatisch konnotiert. Eine manuelle Annotation der Zielhypothesen wurde ebenfalls vorgenommen, während in einem zweiten Schritt auch die Zielhypothesen automatisch mit POS-, lemma (und Morph-Annotationen) versehen wurden. Wichtig für das Lernkorpus sind vor allen Dingen auch die Fehler der Lernenden, die ausführlich annotiert sind. Dadurch, dass das Korpus über eine sogenannte multi-layer-Architektur verfügt ist das Korpus beliebig erweiterbar.

Doch wie verwendet man eigentlich das Falko-Korpus? Dazu folgt nun eine genauere

### **Arbeitsanleitung:**

Für die Suche im Falko Korpus steht das ANNIS2 Interface zur Verfügung.

Es gibt mehrere Fenster/Tools, die beiden Wichtigsten sind das Suchformular (Searchform, welches sich nochmal unterteilt) und das Ergebnisfenster.

In der Mitte werden die (grade) verfügbaren Korpora gezeigt. Durch Klicken auf die Boxen kann ausgewählt werden, in welchem Korpus gesucht werden soll. Wenn man einen bestimmten Korpus nicht findet, kann man auf „more corpora“ klicken, um weitere Korpora aus einem Fenster zu ziehen.

In das „AnnisQL“ werden die Suchanfragen eingegeben. Nachdem man auf den „Show Result“ Button gedrückt hat werden die jeweiligen Positionen der Eingabe in den ausgewählten Korpora angezeigt. Mit den „Context Left/Right“ Buttons lassen sich bis zu zehn weitere Stellen neben dem Resultat anzeigen.

Das Ergebnisfenster zeigt dann zehn gefundene Stellen pro Seite an, zwischen denen man mittels einer Toolbar navigieren kann. Durch die „Citation URL“ Funktion kann man einen Hyperlink mit den Ergebnissen generieren.

Die eingegebenen Begriffe/Tokens werden im Ergebnisfenster rot angezeigt, während die weiteren Satztoken in schwarz gehalten sind. Unter dem jeweiligen Ergebnis lassen sich weitere Annotationsebenen durch das Pluszeichen öffnen.

### **Quellen:**

Reznicek, Lüdeling, Krummes, Schwantuschke: *Das Falko Handbuch: Korpusaufbau und Annotationen*. Hrsg: Humboldt-Universität zu Berlin. Berlin. 2012.

Siemen, Lüdeling, Müller: *Falko – Ein fehlerannotiertes Lernkorpus des Deutschen*. Hrsg: Humboldt-Universität zu Berlin. Berlin. 2012.

[http://www.youtube.com/watch?feature=player\\_embedded&v=TfITTwRDpmg](http://www.youtube.com/watch?feature=player_embedded&v=TfITTwRDpmg)