

SyHD – ein dialekt syntaktisches Korpus und seine Nutzungsmöglichkeiten

Thomas Strobel und Helmut Weiß (Goethe-Universität Frankfurt am Main)

Die empirische Fundierung theoretischer Fragestellungen etwa der (Dialekt-)Syntax gewinnt zunehmend an Bedeutung. Dabei stellt das im Rahmen des DFG-Projekts *Syntax hessischer Dialekte* (SyHD, www.syhd.info, Laufzeit: 2010–2016) gewonnene Datenmaterial zu den im Bundesland Hessen gesprochenen Mundarten – neben vor allem westmitteldeutschen auch niederdeutsche Dialekte sowie Übergänge zu ostmitteldeutschen und oberdeutschen Dialekt-räumen – ein umfangreiches Korpus an indirekt (insgesamt 111 Aufgaben, verteilt auf vier Fragebögen, mit einer Beteiligung von über 1.000 Gewährspersonen und einem Rücklauf von gut 3.500 Fragebögen) wie direkt erhobenen Daten (mehrteilige Interviews an insgesamt 140 Erhebungsorten) dar. Sämtliche Daten liegen glossiert und zum Teil annotiert in einer extra dafür geschaffenen Datenbank vor. Diese erlaubt nicht nur Standardabfragen zu den erhobenen Phänomenen (aus den Bereichen Verbalsyntax, (Pro-)Nominalsyntax, Kongruenz, Wortstellung und Satzverknüpfung), deren Ergebnisse online in Form des dialekt syntaktischen Sprachatlasses *SyHD-atlas* publiziert sind, sondern sie bietet darüber hinaus auch dynamische Tools, mit denen beispielsweise Sekundärauswertungen und Kombinationsabfragen möglich sind. Dies soll im Vortrag anhand ausgewählter Beispiele gezeigt werden. Die verfügbaren dynamischen Komponenten umfassen neben einem Abfragesystem (*SyHD-query*), in dem sich individuell (logische) Abfragen anhand der gesammelten Daten aufbauen und damit eigene Fragestellungen umsetzen lassen, auch ein Auswertungssystem (*SyHD-stats*), das die Daten in Form von Tabellen und Diagrammen zur Verfügung stellt, sowie ein Präsentationssystem (*SyHD-maps*), mit dem interaktiv und individuell Sprachkarten auf der Basis der Projektdaten erstellt werden können. Die vielseitigen Nutzungsmöglichkeiten dieser Online-Werkzeuge und das Potenzial des Korpus, insbesondere auch Korrelationen und Implikationen in den Daten bzw. zwischen den Einzelphänomenen zu finden, sollen anhand von Fallstudien bzw. konkreten Hypothesen vorgestellt werden. Vorrangig werden wir drei Nutzungsmöglichkeiten präsentieren:

- **Rekonstruktion diachroner Entwicklungen anhand synchroner Variation**
Eine erste Nutzungsmöglichkeit ist, dass man mit dem Korpus potenziell den gesamten synchronen Variationsraum einer Konstruktion abdecken und auch seltene Varianten entdecken kann, was wiederum für die Rekonstruktion einer diachronen Entwicklung einen wichtigen Baustein liefert. Ein Beispiel dafür ist der Pronomenzyklus, für dessen Rekonstruktion die in den SyHD-Daten vorhandene Variante einer unmarkierten Abfolge Objektpronomen vor Subjektpronomen Evidenz lieferte.
- **Deduktiv: Überprüfung bestehender Hypothesen**
Eine weitere Nutzungsmöglichkeit ist die Überprüfung von Forschungshypothesen. So eignen sich die SyHD-Daten zum Beispiel dazu, diverse Annahmen zur langen *w*-Extraktion im Deutschen (Subjekt-Objekt-Asymmetrie, Nord-Süd-Asymmetrie, grammatik- vs. parsing-bedingt) zu überprüfen. Dabei bestätigt sich unter anderem die Subjekt-Objekt-Asymmetrie, während sich innerhalb des hessischen Dialektraums kein Nord-Süd-Gefälle feststellen lässt.
- **Induktiv: Gewinnung neuer Hypothesen aufgrund entdeckter Korrelationen**
Ferner konnten mit den Projektdaten nicht nur deduktiv Hypothesen zu den Kookkurrenz-möglichkeiten der Partitivpronomina *ere* (PL und F.SG) und *sen* (M./N.SG) mit Numeralien/Quantoren als Residuum, gestrandeten Adjektiven oder Adjunkten vs. Komplementen (Sätze und PPs) überprüft werden, sondern auch induktiv ein Zusammenhang zwischen dem Partitivpronomen *ere* und „flektierten“ Numeralien (Schwa am Zahlwort) bei nominaler Ellipse gefunden werden, da diese beiden Strategien in den Dialekten Hessens eine deutliche komplementäre areale Distribution aufweisen.