



Transkriptionskonventionen

Entwicklung eigener Transkriptionskonventionen

- Möglichkeit zur Darstellung aller im Ruhrdeutschen auftretenden phonologischen Phänomene
- Basierend auf HIAT (Rehbein & Schmidt 2004), GAT (Selting 2009) und DIDA (Schütte 2004)

Transkription in EXMARaLDA (Schmidt & Wörner 2009)

- EXMARaLDA bietet die Möglichkeit der Kombination von Primärdaten, Metadaten und Transkription.
- Die Verwendung standardisierter Datenformate (z.B. XML, UTF-8, Annotation-Graphs) gewährleistet die Nachhaltigkeit der Daten und ermöglicht die einfache Konvertierung in andere Formate, z.B. für die Präsentation als Webkorpus.
- Eigene virtuelle Tastatur in EXMARaLDA für das KgSR-Projekt



Segmentierung in Intonationseinheiten

- Erstellung einer segmentierten Transkription auf Basis der Prosodie und Pausenstruktur
- Das Zeichen ° markiert das Ende von Intonationseinheiten

Zwei Transkriptionsspuren

- Enge Transkription: Literarische Umschrift

- Verwendung der Standardorthographie
- Abweichungen von der Standardlautung oder typische „Merkmale des Ruhrdeutschen“ (Mihm 1997) werden mit standardorthographischen Mitteln und speziellen Sonderzeichen dargestellt.
- Annäherung an eine phonetische Transkription bei lesbarer Darstellung des Ruhrdeutschen
- Möglichkeit der gezielten Suche nach typischen Phänomenen der gesprochenen Sprache des Ruhrgebiets

- Weite Transkription: Normalisierte Umschrift

- Verwendung der Standardorthographie (inklusive Groß- und Kleinschreibung)
- Morphologische Normalisierung bei sprachlichen Ausdrücken, die im Standard als ungrammatisch gelten:

(1) vor vier Jahre
vor vier Jahren

- Basis für die (semi-)automatische Annotation (z.B. Wortarten-Tagging, Lemmatisierung, syntaktische Analyse)
- Die weite Transkription ermöglicht die einfachere Suche nach Varianten bestimmter Wörter oder Konstruktionen anhand der Standardorthographie.

Transkription phonologischer Besonderheiten

Markierung von abweichender Vokallänge

: für Länge, . für Kürze (vgl. IPA)

(2) we:ch täch schön
Weg Tag schon

Markierung von abweichender Vokalqualität

(3) spete^r äh^rlich
später ehrlich

Konsistente Markierung von Spirantisierung

(4) folchlich könich
folgich König

Sonorisierung

in Anlehnung an IPA markiert mit ,

(5) iṣ=et giḃ=et schlafanzuch
ist es gibt es Schlafanzug

Darstellung der <r>-Vokalisierung

- Die phonetische Transkription der Vokalqualität bei der <r>-Vokalisierung wäre sehr aufwändig und potentiell ungenau.
- Daher wird nur eine phonologische Markierung der <r>-Vokalisierung vorgenommen.

(6) du^rch ga:^rten kinde^r
durch Garten Kinder

Markierung silbischer Konsonanten

mit einem vertikalen Strich | (vgl. IPA)

(7) ihnn[|] heizungn[|]schwimm[|]
ihnen Heizungen schwimmen

Kontraktion und Klitisierungen

Darstellung mit dem in Glossen üblichen Zeichen =

(8) dṑ=nṑ=mḃ da=s au=m schḡ=mḃ

| | 25:02 | 25:03 | 25:04 | 25:05 | 25:06 | 25:07 | 25:08 | 25:09 | 25:10 | 25:11 |
|-----------|---|-------|-----------|---|---|-------|--|--------------------------------|---------------------------------------|-------|
| S1 [engl] | i=mein ich hab wohl büche [°] un sowat alle [°] | | wenn ich | wenn ich i'genwie ma sowat [°] | erzähle ode ^r (so) dann schreib ich mi ^r die ja dinga dann auf [°] | dann | kann ich die ja imme ^r ablesen [°] | so aus=m kopp [°] | ich bin ja kein student [°] | |
| S1 [weil] | Ich meine, ich habe wohl Bücher und sowas alles. [°] | | Wenn ich, | wenn ich irgendwie mal sowas [°] | erzähle oder so, dann schreib ich mir die ja Dinger dann auf. [°] | dann | kann ich die ja immer ablesen. [°] | so aus dem Kopf - [°] | ich bin ja kein Student. [°] | |