FOLKER

Effizientes Transkribieren von Gesprächen

FOLKER wurde für den Einsatz im Projekt „Forschungs- und Lehrkorpus Gesprochenes Deutsch“ (FOLK) optimiert und ist daher auf das effiziente Erstellen von großen Mengen qualitativ hochwertiger und datentechnisch gut zu verarbeitender GAT2-Minimaltranskripte ausgelegt.

Im Gegensatz zu in der Gesprächsforschung verbreiteten Mehrzweck-Tools wie dem EXMARaLDA-Partitur-Editor, ELAN oder Praat versucht FOLKER, durch eine bewusste Reduktion und Fokussierung der Funktionalität auf ein spezifisches Nutzungsszenario die Erlernbarkeit und Bedienbarkeit zu erleichtern.

Transkribieren in FOLKER

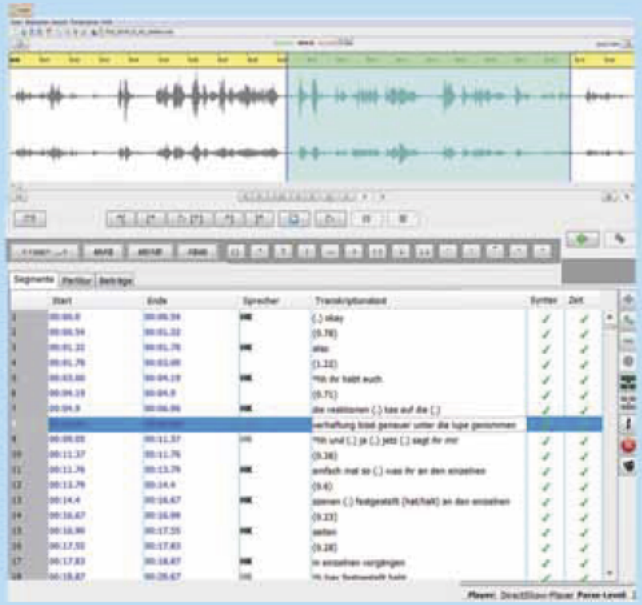
Neben einer Darstellung des Audiosignals in Form eines **Oszillogramms** bietet FOLKER drei verschiedene, mit dem Audio synchroniserte Ansichten des Transkripts, die auf unterschiedliche Schritte im Erstellungsprozess abgestellt sind.

Die **Segment-Ansicht** eignet sich besonders für die Ersteingabe – sie erlaubt eine präzise Zuordnung zwischen kurzen Abschnitten aus der Aufnahme und zugehörigen Textsegmenten.

Die **Partitur-Ansicht** ist vor allem für die präzise Transkription von Simultanpassagen geeignet. In der **Beitrags-Ansicht** werden zusammenhängende Segmente eines Sprechers in größeren Einheiten zusammengefasst, was ein abschließendes Korrekturlesen und -hören des Transkripts erleichtert.

Während der Eingabe kann der Transkriptionstext auf **Konformität** mit den Vorgaben für das GAT-Minimal- oder Basistranskript geprüft werden. Ein unzulässiger Einsatz von Transkriptionszeichen wird dem Nutzer dann unmittelbar signalisiert.

Der in FOLKER eingebaute **Player** erlaubt ein präzises und flexibles Navigieren in der Audio-Datei in Abhängigkeit von der momentanen Position im Transkript.



Ausschnitt aus dem FOLKER-Nutzerinterface mit Oszillogramm und Transkript in Segmente-Ansicht

Transkript in Partitur-Ansicht

```
<contribution speaker-reference="DRE3"
    start-reference="TLI_2268" end-reference="TLI_2270"
    parse-level="2">
    <w>ja</w>
    <pause duration="micro"/>
    <w>ich</w>
    <w>bin</w>
    <w>befreit</w>
    <non-phonological description="lacht"/>
</contribution>
```

Transkript als XML-Datei

Die Oszillogramm-Ansicht erleichtert dabei das Festlegen geeigneter Segmentgrenzen.

Weitere Funktionen unterstützen eine systematische Maskierung von Transkript und Audio-Datei mittels eines Maskierungsschlüssels.

Interoperabilität

FOLKER speichert Transkripte in einem **XML-Format**, in dem alle relevanten Einheiten (Sprecher, Segmente, Wörter, Pausen, Beiträge etc.) in strukturierter, maschinenlesbarer Form repräsentiert werden. FOLKER-Dateien können somit zum Beispiel ohne weitere manuelle Bearbeitungsschritte in die DGD2 übernommen werden.

Da FOLKER auf dem **EXMARaLDA**-Datenmodell basiert und mit diesem vollständig kompatibel ist, lassen sich auch die Tools des EXMARaLDA-Systems nutzen, um FOLKER-Daten weiter zu bearbeiten. Weitere Import- und Exportmöglichkeiten existieren für den Austausch mit den Programmen **ELAN**, **Praat** und **F5.**

Für die Ausgabe von Transkripten auf Bildschirm oder Papier stellt FOLKER weitere Exportmöglichkeiten im **HTML**- oder **RTF**-Format zur Verfügung.

OrthoNormal

Als Ergänzung zu FOLKER wurde in FOLK das Tool OrthoNormal entwickelt, das ein effizientes **orthographisches Normalisieren** von Transkripten in literarischer Umschrift unterstützt. Darüber hinaus ermöglicht OrthoNormal das manuelle Korrigieren der Resultate einer automatischen Lemmatisierung und eines POS-Taggings..

Download und Support

FOLKER und OrthoNormal werden für die Betriebssysteme Windows, MAC OS X und Linux kostenlos zum Download angeboten **(http://agd.ids-mannheim.de/folker.shtml)**. Mitarbeiter des AGD bieten Unterstützung und Beratung bei der Verwendung der Tools. In regelmäßigen Abständen finden am IDS Schulungen für FOLKER-Nutzer statt.

Forschungsgegenstand

FOLKER – der FOLK EditoR – ist eine Software für die Transkription von Audioaufnahmen.
FOLKER wird als Transkriptionstool für das Forschungs- und Lehrkorpus Gesprochenes Deutsch (FOLK) verwendet und im Projekt kontinuierlich weiterentwickelt. Er stellt eine Adaption des EXMARaLDA-Systems an die Bedarfe des FOLK-Projekts dar.

FOLKER ermöglicht ein effizientes und präzises Transkribieren von Gesprächsdaten nach den Konventionen des Gesprächsanalytischen Transkriptionssystems (GAT2).

An FOLK angegliederte Tools unterstützen die orthographische Normalisierung von Transkripten in literarischer Umschrift, das Lemmatisieren und POS-Tagging von Transkriptionsdaten, sowie ein Maskieren von Transkripten und Audio-Aufnahmen. FOLK wird kostenlos zum Download angeboten.

Dr. Thomas Schmidt
Leiter des Programmbereichs Mündliche Korpora

Projekt-Kontakt

Dr. Schmidt, Programmbereich Mündliche Korpora
Institut für Deutsche Sprache, Mannheim
agd@ids-mannheim.de

Beteiligte Projekte

- Forschungs- und Lehrkorpus Gesprochenes Deutsch
- EXMARaLDA

Beteiligte Mitarbeiterinnen und Mitarbeiter

- Dr. Thomas Schmidt
- Dr. Wilfried Schütte
- Jenny Winterscheid

INSTITUT FÜR
DEUTSCHE SPRACHE

R 5, 6-13 • D-68161 Mannheim
Tel.: +49 621 1581-0
Fax: +49 621 1581-200
info@ids-mannheim.de
www.ids-mannheim.de

Das Institut für Deutsche Sprache (IDS) ist die zentrale Einrichtung zur Erforschung und Dokumentation der deutschen Sprache in ihrem Gebrauch und in ihrer neueren Geschichte. Es gehört zu den 86 außeruniversitären Forschungs- und Serviceeinrichtungen der Leibniz-Gemeinschaft.

50 JAHRE INSTITUT FÜR DEUTSCHE SPRACHE

Abteilung Pragmatik

Institut für Deutsche Sprache

Stand 01/2014

FOLKER

Ein Transkriptionseditor für das effiziente Transkribieren von Gesprächen

Mitglied der Leibniz-Gemeinschaft