

Csaba Földes

Ungarndeutsche Sprachvariation und Mehrsprachigkeit

Ein Korpusprojekt auf der Basis von empirischer Feldforschung und Online-Sprachdokumentation

Abstract

The paper at hand presents intentions, contents and in particular operational procedures regarding a research project in progress from September 1st 2015 to December 31st 2017. The project in question is called "Digital portal 'Bilingual and language contact corpus of German as a minority language in Hungary'" and is located at the Chair of Germanic Linguistics at the University of Erfurt/Germany. Its main goals are research and documentation: Interactive practices and their linguistic correlates are to be discovered based on German as a minority language in Hungary within an inter- or rather transcultural context of bilingualism. The linguistic corpus compiled within the framework of this project will be accessible online in form of a structured and commented electronic database. Thus, the corpus will present authentic spoken linguistic material from various important regions of settlement of German minorities. Presenting a work in progress, this article intends in this context to contribute up to date information on a specific area of current contact-driven German-Hungarian bi- or rather multilingualism somewhere between qualitative field studies and online language documentation.

Keywords: German as a minority language in Hungary, bilingualism, languages in contact, language documentation

1 Präliminarien

Das Beziehungsfeld Deutsch-Ungarisch stellt einen etablierten sprachwissenschaftlichen Untersuchungsgegenstand mit breitem Erkenntnishorizont sowohl im Hinblick auf die Sprachvergleichs- als auch bezüglich der Sprachkontaktdimension dar; vgl. für das erstere z.B. die

beeindruckende kontrastive Grammatik von Pilarský (2013) in drei Bänden, für das zweite z.B. die voluminöse Monographie von Mollay (1982). Die synchrone Kontaktproblematik scheint hingegen – trotz einiger bemerkenswerter Arbeiten – eine weniger prominente Rolle im Forschungsdiskurs zu spielen. Vor diesem Hintergrund beabsichtigt der vorliegende Aufsatz, Anliegen, Inhalte und Verfahren eines derzeit laufenden Forschungsvorhabens zu thematisieren, um dadurch zu einem spezifischen Bereich der gegenwärtigen kontaktgeprägten deutsch-ungarischen Zwei- bzw. Mehrsprachigkeit auf der Basis von Feldforschung und Online-Sprachdokumentation einen aktuellen Beitrag zu leisten. Im Betrachtungsfokus steht also das von der Beauftragten der Bundesregierung für Kultur und Medien (BKM) geförderte Projekt "Digitales Portal 'Ungarndeutsches Zweisprachigkeits- und Sprachkontaktkorpus'" (UZSK).¹ Zeitraum: 01.09.2015–31.12.2017, Forschungsstelle: Lehrstuhl für Germanistische Sprachwissenschaft an der Philosophischen Fakultät der Universität Erfurt.

2 Inhalts- und Zielstruktur des Projekts

Der Gegenstand ist eine Mehrsprachigkeitskultur mit spezifischen Ausprägungsstrukturen und Verwendungsmustern des Deutschen. Das Vorhaben ist disziplinär zum einen im Rahmen der diatopisch orientierten Variationslinguistik (Stichwort Regionalsprachen- bzw. Sprachdynamikforschung), zum anderen in dem der Zwei- bzw. Mehrsprachigkeitsforschung und der Kontaktlinguistik anzusiedeln. Für seine Durchführung kommt ergänzend die Korpuslinguistik hinzu.

Der Inhalt des Vorhabens liegt – im Sinne einer Doppelperspektive – in der Untersuchung und der Dokumentation² zweisprachiger Sprechhandlungen im sog. bilingualen Diskursmodus am Material des Deutschen als Minderheitensprache in Ungarn (Stichwort: "Donauschwaben"). Daher richtet sich das Projekt zum einen auf die variations- und kontaktlinguistische Erforschung ungarndeutscher mündlicher Sprechhandlungen im interaktiven Alltag der Verständigung

¹ Nummer des Projekts: K 44-41000/1#92. Für die Bewilligung und die Förderung sei an dieser Stelle unser Dank ausgesprochen.

² Damit wird auch dem Postulat z.B. von Chelliah (2013: 69f.) entsprochen, dass reine Dokumentation auch ohne Deskription möglich sei, eine Verbindung von beiden jedoch "ideal" ist.

und zum anderen auf die Erstellung eines webbasierten Portals inklusive Datenbank für authentische ungarndeutsche Diskurse in Form von Tonaufnahmen, Transkripten und Texten.

Der Hintergrund: Während sich das Gros der bisherigen Forschungen auf eine traditionelle – z.B. dialektgeographische – Beschreibung der ungarndeutschen Basismundarten (Bairisch, Fränkisch und Schwäbisch) konzentriert hat, stellt dieses Forschungsvorhaben die durch Mehrsprachigkeit, Inter-/Transkulturalität und (interne sowie externe) Variation weitgehend geprägte tatsächliche Sprachverwendung in den Mittelpunkt. Denn die "reinen" (d.h. ungemischten) deutschen Dialekte werden in der Untersuchungsregion heute kaum noch gesprochen, vielmehr sind dynamische Misch- und Übergangsformen mit Ungarisch (und seltener mit anderen Kontaktsprachen) charakteristisch. Sind doch gleichsam alle Ungarndeutschen zwei- bzw. mehrsprachig mit unterschiedlichen Dominanzkonfigurationen, jedoch überall mit einem deutlichen Übergewicht der ungarischen Standardsprache.³ Das als Ergebnis vorgesehene ungarndeutsche Zweisprachigkeitskorpus soll reale, aus dem Leben gegriffene Gespräche systematisch aufbereiten, die die ungarndeutschen Kontaktvarietäten der Gegenwart in ihrer Vielfalt anschaulich repräsentieren. Nach Himmelmann (1998: 185f.) kann man grundsätzlich vier Typen von Sprechereignissen unterscheiden: (a) natürliche, (b) beobachtete, (c) evozierte kommunikative Ereignisse und (d) Elizitationen. Im Blick des Projekts stehen die Typen (b) und (c).

Das Hauptziel besteht mithin in Forschung und Dokumentation: Interaktive Praktiken und ihre sprachlichen Korrelate sollen in einem inter- bzw. transkulturellen Zweisprachigkeitskontext aufgedeckt werden. Das im Rahmen des Projekts erarbeitete Korpus soll – in strukturierter und kommentierter Form – in einer elektronischen Datenbank digital zugänglich und nutzbar gemacht werden. Dementsprechend wird das UZSK authentisches sprechsprachliches Material

³ Gewöhnlich resultiert die Sprachkompetenz ungarndeutscher Sprecher vor allem aus dem jeweiligen deutschen Ortsdialekt, aus der ungarischen Standardsprache und meist in geringerem Maße aus der deutschen Standardsprache. Von einer weiteren Darstellung von Sprache und Kommunikationsprofil der Ungarndeutschen soll an dieser Stelle abgesehen werden; zu den ungarndeutschen Dialekten und ihrer Erforschung vgl. z.B. den Übersichtsartikel von Manherz (2014), zur soziolinguistischen Ausgangslage und zur sprachlichen Binnen- und Kontaktstruktur findet man ausführliche Beschreibungen z.B. in den Publikationen von Erb (2010: 120 ff.), Földes (2013: 121 ff.) und Knipf-Komlósi (2011: 17 ff.).

aus mehreren wichtigen Siedlungsregionen deutscher Minderheiten bereitstellen. Dabei gilt den zahlreichen und vielgestaltigen Sprach- und Kulturkontaktphänomenen aus dem Ungarischen (gelegentlich aus anderen Minderheitensprachen), die heute ein hervorstechendes Merkmal ungarndeutscher Redeweise verkörpern, besonderes Augenmerk. Somit soll unser Korpus über seine ungarndeutsche Relevanz hinaus allgemeiner für Zweisprachigkeits- und kontaktlinguistische Studien auch in Bezug auf andere Sprachkonstellationen von Interesse sein.

3 Projektdetails und Ausführungsschritte

3.1 Zu Grundkonzept, Hauptlinien und zentralen Positionen dieses Forschungs- und Dokumentationsvorhabens wurde im Hinblick auf dessen Anlaufphase bereits in zwei Publikationen referiert (Földes 2016a und 2016b), so dass sich dieser Artikel im Folgenden auf den derzeitigen Stand der Arbeiten konzentriert, in der Hoffnung, dass dieser inhaltliche Werkstattbericht in gewissem Maße auch allgemeineren Aufschluss bietet. Im Mittelpunkt der Betrachtungen steht die Dokumentationskomponente, vor allem die Funktionalität der Datenbank; dabei wird hier auf die linguistische Forschungskomponente nicht näher eingegangen.⁴

3.2 Der hohe und rasche Informationsbedarf unserer Zeit verlangt einen schnellen und zuverlässigen Informationszugang, der dank der allgemeinen Verbreitung der Computer- und Software-Technik heutzutage immer mehr zu einer Selbstverständlichkeit wird. Dieser Anspruch färbt grundlegend auf die gegenwärtige Linguistik ab; aus den hervorragenden technologischen Möglichkeiten ergeben sich innovative analytische Perspektiven. So steht nunmehr eine breite Palette unterschiedlichster Sprachkorpora zur Verfügung. Korpuslinguistische Arbeiten haben das Kulturphänomen Sprache in seinen verschiedenen Erscheinungsformen zum Gegenstand, wobei sich das Interesse auf immer mehr – z.T. spezifische – Varietäten und Sonderformen erstreckt. Die Korpuslinguistik ist dabei auf authentische (mündliche und schriftliche) Sprachdaten angewiesen, die zu größeren Textkor-

⁴ Frau Uschi Schmidt, M.A., wissenschaftliche Mitarbeiterin im Rahmen des Projekts, danke ich für ihre wertvolle Mit- und Zuarbeit.

pora zusammengeführt werden sollen. Angesichts der Fülle der bereits bestehenden Korpora überrascht zunächst, dass Zweisprachigkeitskorpora kaum vorliegen: Zweisprachige Korpora gibt es zwar in hoher Zahl, aber solche, die zweisprachiges Sprechen, also den Diskurs von zweisprachigen Personen, dokumentieren, sind eine absolute Rarität.

Unser Hauptanliegen – Erarbeitung eines Online-Portals mit einem ungarndeutschen Zweisprachigkeits- und Sprachkontaktkorpus – setzt theoretische Konzepte, operationale Techniken und empirische Datenerträge von Feldforschung und Sprachdokumentation voraus. Diesbezüglich musste sich das Projektdesign großen Herausforderungen stellen. Denn beispielsweise hat Tagliamonte (2006: 17) ironisch festgestellt: "Feldforschungsmethoden sind wahrscheinlich das am besten gehütete Geheimnis der Soziolinguistik". Außerdem ist Schilling (2013: 1) zuzustimmen, wenn sie beklagt, dass Feldmethoden, bei denen linguistische Daten gewonnen werden, unscheinbar sind; die meisten Wissenschaftler konzentrieren sich auf Ergebnisse und Interpretationen, ohne zu fragen, woher die Daten stammen. Feldforschung heißt eigentlich, empirisch sprachbezogene Daten in der natürlichen Umgebung zu sammeln (Bowerman 2015: 2). Genau das ist der erste Schritt unseres Vorhabens.

3.3 Auch die in der Regel als Oberbegriff fungierende "Sprachdokumentation" gehört, zumal im Rahmen der Sprachkontaktforschung, nicht zu den besonders frequenten Themen der Gegenwartslinguistik. Gleichwohl erscheinen hin und wieder ambitionöse Publikationen, die anhand ihres Titels viel Einschlägiges vermuten lassen. Zum Beispiel kürzlich der Band von Comrie & Golluscio (2015) "Language Contact and Documentation", der eine Auswahl von Beiträgen vereint, die auf dem "II International Symposium on Language and Culture Documentation in Latin America in Buenos Aires" in Argentinien vorgestellt wurden. Der Herausgeber und die Herausgeberin bezeichnen diese Aufsätze als äußerst relevant für die aktuelle regionale und internationale Diskussion um bedrohte Sprachen, Sprachkontakt, Dokumentation und Arealtypologie, wobei der Fokus auf südamerikanischen Sprachen liegt (S. 1). Darüber hinaus beinhaltet das Buch Beiträge, die neue Einsichten in die Dokumentation sowohl bedrohter als auch "kleiner" Sprachen und Sprachkontaktsituationen in weiteren Gebieten rund um den Globus gewähren, wie im Mittleren Osten, Australien und Papua-Neuguinea. Zu den übergeordneten

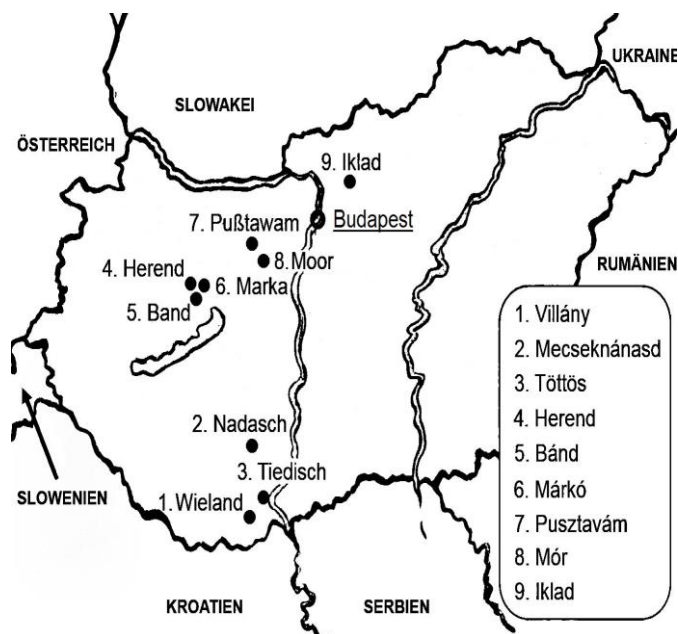
Themen gehören Fallstudien in Dokumentation, technischer Umsetzung sowie Forschung und Bewahrung der Sprachen der Welt. Spezifischer geht es unter anderem um mündliche Traditionen und kulturelle Bedeutungen von Multilingualität, um Beziehungen von Sprachen untereinander und die daraus folgenden Konsequenzen wie Mischung, Erwerb oder Verlust, Sprachkontakt durch geographische Nähe etc. – unter regionalen und kulturellen Bedingungen, die von den ungarndeutschen Settings wesentlich abweichen.

Bei vielen Leistungen des Bandes ist jedoch der im Titel verzeichnete Dokumentationscharakter in manchen seiner Beiträge unklar bzw. nicht explizit genug; einige Texte, wie z.B. Muysken, Hammarström, Birchall, van Gijn, Krasnoukhova & Müller (2015), setzen sich gar nicht mit Dokumentation auseinander. Für die Belange des in Frage stehenden ungarndeutschen Forschungsvorhabens bietet das Buch wenig Aufschluss.

3.4 Wie auch Schilling (2013: 3) betont, werden die besten Reflexionen der Feldmethoden aus Projekten gezogen, bei denen mindestens ein gewisses Maß an Insider-Involvierung besteht. In diesem Sinne wird das Material – in unserem Auftrag und aufgrund unserer Richtlinien – von ortsansässigen und in der Regel auch dialektkundigen ungarndeutschen Germanist(inn)en (mittels gelenkter Gespräche bzw. halbstandardisierter Interviews) erhoben, die einen direkten Zugriff auf die Sprecher(innen) haben und als Angehörige der Minderheit bzw. der Ortsgemeinschaft deren Vertrauen weitgehend genießen.

Zur Erreichung einer größtmöglichen Repräsentativität werden von der Region und vom Dialekt her verschiedene Orte berücksichtigt. Grafik 1 situiert die betreffenden Ortschaften innerhalb Ungarns, wobei das Landkartenbild die deutschen, die Liste unten rechts die ungarischen Ortsnamensvarianten zeigt, also deutsch: Wieland, ungarisch: Villány, Nadasch/Mecseknádasd, Tiedisch/Töttös, Herend/Herend, Band/Bánd, Marka/Márkó, Pußtawam/Pusztavám, Moor/Mór und Iklad/Iklad. Man kann also erkennen, dass auf Transdanubien als wichtigste Siedlungsgegend der Ungarndeutschen fokussiert wird mit den Schwerpunkten Ungarisches Mittelgebirge im Norden (eine vorwiegend durch bairische *ua*-Mundarten geprägte Sprachlandschaft) und "Schwäbische Türkei" im Süden (ein Areal eher fränkischer Prägung). Außerdem ist Iklad, nordöstlich von Budapest gelegen, einbezogen worden, nicht zuletzt, da es eine multiethni-

sche und dreisprachige Gemeinde ist, in der neben Ungarisch und Deutsch traditionell auch Slowakisch eine Rolle spielt, so dass in diesem Material auch mit multiplen Sprachkontaktphänomenen zu rechnen ist.



Grafik 1: Die Erhebungsorte

Die typologische Einordnung des UZSK ließe sich – unter Rückgriff auf die Gliederung von Scherer (2014: 17-32) – wie folgt vornehmen (vgl. auch Földes 2016b). (1) Aufgrund formaler Kriterien: (a) nach dem Speichermedium: computerlesbares Korpus, (b) nach der Hierarchie: Teilkorpus, (c) nach der Vollständigkeit der Texte: Volltextkorpus, (d) nach der Abgeschlossenheit: statisches Korpus,⁵ (e) nach der Aufbereitung des Korpus: (zumindest teil)annotiertes Korpus, (f) nach dem Medium:⁶ unimodales Korpus;⁷ (2) aufgrund inhaltlicher

⁵ Gleichwohl könnten im Falle einer Verlängerung der Projektdauer weitere ungarndeutsche Siedlungsgebiete mit einbezogen werden, wodurch eine quantitative Erweiterung der sprachlichen Datenbasis möglich wäre.

⁶ Bei Scherer (2014: 17) gibt es diesen Gliederungsaspekt nicht, er ist aber in der Literatur gängig (beispielsweise bei Mello 2014: 29).

Kriterien: (a) nach dem Sprachmedium: Korpus der gesprochenen Sprache, (b) nach dem zeitlichen Bezug: Korpus der Gegenwartssprache, (c) nach dem Geltungsbereich des Korpus: Spezialkorpus und (d) nach der Anzahl der Sprachen: einsprachiges Korpus.⁸ Nach der Terminologie von McEnery & Hardie (2012: 11) kommt unser Korpus, wie wir es im momentanen Stadium einschätzen, in die Nähe des Typs "opportunistic corpus".

3.5 Die Audiotranskription der ermittelten sprachkommunikativen Daten erwies sich für das Projekt als ein überaus schwieriges Aufgabenfeld. Das Arsenal der (manuellen wie elektronischen) Transkriptionskonzepte und -systeme ist mittlerweile beträchtlich: Die Linguistik und die qualitative Sozialforschung bieten nunmehr eine Bandbreite verschiedener Konventionen. Bekanntlich hängt dabei die Auswahl des eingesetzten Transkriptionsverfahrens u.a. von Forschungsziel und -methodik, Erkenntniserwartung, intendierter Verwendung und auch von forschungspragmatischen Gründen in der konkreten Situation ab (vgl. Dresing & Pehl 2015: 18 und 20). Phonetische Feinheiten (artikulatorische Nuancen etc.) und prosodische Details (Tonhöhenverlauf, Nebenakzent, Lautstärke, Sprechgeschwindigkeit etc.), die ein enges Feintranskript bieten könnte, sind z.B. in unserem Forschungsrahmen nicht ausschlaggebend; uns kommt es prinzipiell auf die Dokumentation und Verdeutlichung bilingualer Sprechproduktion an, also darauf, wie sich Varietäten des Deutschen und des Ungarischen in den Diskursen zweisprachiger Ungarndeutscher manifestieren und miteinander interagieren.

Es fällt auf, dass trotz der Vielzahl der in der Forschungsliteratur vorkommenden Umschriften ein für die vorliegenden Projektzwecke völlig geeigneter Transkriptionsmodus noch aussteht. Es fehlt also ein adäquates System, das Aspekten des Diskursverhaltens, der Dialektalität und der zwei- bzw. mehrsprachigen Sprachproduktion zugleich Rechnung trägt. Vergleichbare Projekte, die der Sprache deutscher Minderheiten auch unter dem Blickwinkel ihrer Zweisprachig-

⁷ Es handelt sich im Wesentlichen um Unimodalität, weil ja primär Audiodateien präsentiert werden. Da aber viele von ihnen auch schriftlich transkribiert werden, könnte man z.T. von einer Art Multimodalität sprechen.

⁸ Im Namen kommen zwar zwei Sprachen vor, aber die erfassten Diskurse sind im Prinzip immer einsprachig, da die Matrixsprache Deutsch ist, wenngleich dieses Deutsch diverse Kontaktphänomene des Ungarischen aufweist.

keit gewidmet sind, verwenden in der Regel ganz simple (und oft mehr oder weniger improvisierte) Notationen. Zum Beispiel benutzt das umfassende "Texas German Dialect Project"⁹ laut E-Mail-Mitteilung von dessen Leiter, Hans C. Boas, schlicht und einfach eine "modifizierte deutsche Orthographie": IPA oder GAT 2 (Gesprächsanalytisches Transkriptionssystem) wurden nicht eingeführt, weil beide sehr aufwendig seien und viel Zeit und Geld kosten und es sehr lange dauere, bis man gute Leute trainiert habe, die auch so einheitlich transkribieren können. So seien die Texte hinterher auch für Nicht-Linguisten lesbar.¹⁰

In dem sonst informativen Interviewband von Lasatowicz & Pelka (2011) zu oberschlesischen Sprachbiographien werden in der Einführung zwar "interessante Einblicke in die Phonetik [...]" der Gewährspersonen in Aussicht gestellt – "man denke hier z.B. an die von der standardsprachlichen abweichende Realisierung einzelner Vokale [...]" (S. 9), doch mögen gerade diese besonders interessanten Phänomene zwar vielleicht in den Tonaufnahmen enthalten sein, die Transkription im Buch hat diese jedoch völlig verschwinden lassen. Denn: "Das Schriftbild einzelner Lexeme, deren Zuordnung zu einer Sprache schwierig erscheint, wird im Falle von Lauten, die gleich ausgesprochen werden, innerhalb von deutschen Passagen der deutschen Orthographie angepasst, innerhalb von polnischen Passagen – der polnischen" (S. 12). Was noch markanter auffällt: In den einleitenden Passagen zur Transkription wird deklariert, dass GAT verwendet werde (S. 12), währenddessen die abgedruckten Interviews gleichsam eine normale Verschriftlichung bieten (mit Großbuchstaben, Satzzeichen etc.) und mit GAT offenbar nicht viel gemein haben.

Hinsichtlich seiner Anlage scheint das von Nina Berend geleitete und ebenfalls von der BKM geförderte Projekt "Elektronisches Wörterbuch. Ein Online-Informationsangebot zu Sprache und Dialekten der Russlanddeutschen" am Institut für Deutsche Sprache in Mannheim mit seiner "Russlanddeutschen Dialektdatenbank" (RuDiDat) unserem Projekt am nächsten zu stehen. Obwohl dieses Russlanddeutsch-Vorhaben laut Projekt-Flyer (Stand: 09/2015) am 31.12.2015 auslief, ist die Datenbank noch (Ende April 2016) nicht im Netz, so

⁹ Vgl. den Internetauftritt unter <http://tgdp.org/> (Zugriff: 27.04.2016).

¹⁰ Siehe zu den Details das "Texas German Dialect Project Manual" vom November 2003 (Manuskript).

dass wir auf seine Ideenwelt und seine Verfahrenstechniken nicht rekurrieren können.

In ungarndeutschem Bezugsrahmen ist eine sog. literarische Umschrift verbreitet, also eine weite (somit keine lautgetreue) Transkription, die von Ruoff (1973) an der Tübinger Arbeitsstelle für Sprache in Südwestdeutschland entwickelt worden ist. Sie bedient sich der Grapheme und Zeichen der Normalschrift und soll leicht lesbar und handhabbar sein (vgl. auch Knipf-Komlósi 2011: 11).

Angesichts unseres Projektzuschnitts- und -anspruchs erschien uns zu Beginn das Transkriptionsinstrument DIDA, das nicht nur ein Transkriptionsmodell, sondern auch gleichzeitig ein Transkriptionsprogramm ist, am besten zur Veranschaulichung der festzuhaltenden bi- bzw. multilingualen Phänomene geeignet (vgl. dazu Földes 2016b). Entwickelt wurden Programm und Modell in der Abteilung Pragmatik des Instituts für Deutsche Sprache in Mannheim (IDS), von 1992 bis 2004 kamen sie im Rahmen der Verschriftlichung von Sprachaufnahmen zum Aufbau von Korpora gesprochener Sprache zum Einsatz. Eine Weiterentwicklung findet jedoch nicht mehr statt. An Stelle von DIDA ist dort der Transkriptionseditor FOLKER getreten. Daneben existieren auch noch weitere Programme wie zum Beispiel EXMARaLDA.¹¹ Für die Anfertigung von Transkriptionen schweben uns für das Projekt gegenwärtig der Einsatz eines einfachen gesprächsanalytischen Umschriftsystems und die Nutzung eines Programms vor, das die Eingabe nach dem Transkriptionsmodell DIDA erlaubt, eine unkomplizierte Zusammenarbeit zwischen räumlich voneinander getrennten Mitarbeiter(inne)n ermöglicht und die Ausgabe in der Datenbank in eine für die Nutzer(innen) einfache und umfangreich weiterverwendbare Form erlaubt. Diese grundlegenden Anforderungen sollten aktuelle Transkriptionseditoren ohnehin erfüllen. Letztendlich zählt für uns vor allen Dingen die angemessene Dokumentation des Sprachmaterials und folglich die Form seiner Darstellung in der Datenbank.

Daher wird im Projekt schließlich ein Konzept eingesetzt, das im Wesentlichen auf dem Minimaltranskript der Version 2 des "Gesprächsanalytischen Transkriptionssystems" (abgekürzt: GAT)¹² basiert.

¹¹ Detailliertere Informationen findet man auf der Internetpräsenz des Archivs für Gesprochenes Deutsch des IDS unter <http://agd.ids-mannheim.de/dida.shtml> (Zugriff: 25.04.2016).

¹² Siehe das GAT auf der Webseite des Instituts für Deutsche Sprache in Mannheim: <http://agd.ids-mannheim.de/gat.shtml> (Zugriff: 25.04.2016).

Zu GAT 2 als Ausgangspunkt ist zusammenfassend Folgendes zu sagen: Seit dem Erscheinen der ersten Fassung im Jahr 1998 hat sich das GAT im deutschsprachigen Raum als Standard im Bereich der empirischen Kommunikationsforschung durchgesetzt. Die überarbeitete Fassung "GAT 2" wurde 2009 veröffentlicht (Hagemann & Henle 2014: 1). Durch GAT sollten einheitliche Transkriptionskonventionen zur Verschriftlichung gesprochener Sprache geschaffen werden, die es ermöglichen, unabhängig und unkompliziert Gesprächsdaten festzuhalten, weiterzugeben und gesprächsanalytisch zu verwenden. Die ursprüngliche Version von GAT bot Konventionen zur Erstellung von Basis- und Feintranskripten, während mit GAT 2 zusätzlich das Minimaltranskript mit weiter vereinfachten Regelungen eingeführt wurde; vgl. <http://agd.ids-mannheim.de/gat.shtml> (Zugriff: 26.04.2016).

Von diesem Minimaltranskript wird im Projekt ausgegangen, wobei wir die Regelungen im Sinne unseres Forschungsinteresses weiter reduziert haben, wie im Folgenden in einer kurzen Gegenüberstellung unserer Transkriptionsrichtlinien mit dem GAT-2-Minimaltranskript genauer erläutert wird.

Beibehaltene Konventionen: Folgende Richtlinien wurden übernommen, da sie die Lesbarkeit und später eine eventuelle Übertragung in Transkriptionseditoren wie FOLKER oder EXMA-RaLDA gewährleisten.

Zu Beginn werden die **Metadaten** des Gesprächs festgehalten: Aufnahmeort, Aufnahmedatum, Geschlecht und Alter oder Geburtsjahr des Explorators, Geschlecht und Alter oder Geburtsjahr der Gewährperson (d.h. des Befragten) sowie ihr Bildungsstand bzw. Beruf und Herkunfts- bzw. Geburtsort.

Verwendet wird die Schriftart Courier New (möglich ist auch eine andere nicht-proportionale Schriftart) in Schriftgröße 12 pt mit einem Zeilenabstand von 1,5. Der Nutzen dabei ist, dass so **simultanes Sprechen** dargestellt werden kann, indem man die Äußerungen direkt über- bzw. untereinander schreibt und mit Leerzeichen an die entsprechenden Stellen rückt.

Es werden keine Formatierungen für Absätze oder Flächen (Tabulator, Tabellen etc.) sowie keine Worttrennung am Zeilenende eingesetzt.

Es wird durchgehend **Kleinschreibung** verwendet; Großbuchstaben sind für Akzentmarkierungen vorgesehen (mehr dazu un-

ter den "abweichenden Richtlinien"), vgl. auch Hagemann & Henle (2014: 7).

Es wird eine Segmentierung vorgenommen, d.h., es findet eine Untergliederung der turns (wechselnde Redebeiträge der verschiedenen Sprecher) in Intonationsphrasen durch Zeilenumbrüche statt. Dies geschieht, da keine Satzzeichen zur Verfügung stehen und damit kein Fließtext entsteht. Simultanes Sprechen wird wie oben erwähnt dargestellt, indem die jeweiligen Äußerungen in Zeilen untereinander und mit Hilfe von Leerzeichen an die entsprechenden Stellen gesetzt werden.

Die einzelnen Sprecher werden durch Sprechersiglen markiert, zum Beispiel "I" und "B" für Explorator und Gewährsperson oder nummeriert bei mehreren Befragten ("B1", "B2") oder als Sprecher mit "S1", "S2" usw.

Als Notation für annähernd standardsprachlich Realisiertes ist die deutsche Standardorthographie vorgesehen, das heißt: Was relativ standardgemäß ausgesprochen wird, soll nach den Regeln der Orthographie transkribiert werden. Beispiel: Reduktionssilben entsprechen den sprechsprachlichen Standards und werden standardorthographisch notiert ("gehn" als gehen). Nennenswerte phonetische Abweichungen von den Standards werden in ihrer jeweiligen sprechsprachlich realisierten Besonderheit erfasst und mit den Mitteln der Schrift nachmodelliert: z.B. Tilgungen ("n" für *ein* als n; "ne" für *eine* als ne; "un" für *und* als un etc.), Klitisierungen ("biste" für *bist du* als biste; "hamse" für *haben sie* als hamse oder als ham_se etc.) und Regionalismen ("weisch" für *weißt du* als weisch, "Schol" für *Schule* als schol etc.), vgl. auch Selting, Auer & Barth-Weingarten u.a. (2009: 360 ff.).

Des Weiteren werden Komposita ohne Bindestrich transkribiert, z.B. bildzeitung.

Abkürzungen und Buchstabierungen werden ausgeschrieben bzw. in Sprechsilben aufgelöst ("SPD" als es pe de); Zahlwörter werden ebenfalls ausgeschrieben ("22" als zweiundzwanzig oder – je nach Aussprache – zwoundzwanzig), vgl. auch Hagemann & Henle (2014: 8).

Außerdem werden Verzögerungssignale und andere Sprechersignale erfasst (wie z.B. hm, ach, ähm).

Unverständliche Stellen werden in einfache runde Klammern gesetzt: die unverständliche Passage wird durch Leerraum je geschätzter Länge () dargestellt oder mit Angabe der Dauer in dop-

pelten runden Klammern ((unverständlich, 3 Sek.)). Schwer verständliche Stellen werden mit vermuteter Äußerung in einfache runde Klammern gesetzt: (solche) für nicht sicher rekonstruierbaren Wortlaut oder (welche/solche) bei möglichen Alternativen, zwischen denen nicht sicher entschieden werden kann.

Abweichende Richtlinien: Auf einige der im Minimaltranskript von GAT 2 vorhandenen Konventionen wird verzichtet, um nur jene Regelungen beizubehalten, die für unsere Forschungsfragen bzw. Projektziele relevant sind. Daher wird auf die Akzentmarkierung durch Großbuchstaben verzichtet und durchgängig Kleinschreibung verwendet. Das Gleiche gilt für die Markierung von Tonhöhenbewegungen durch Satzzeichen und Gedankenstrich, die ebenfalls nicht verwendet werden, vgl. auch Hagemann & Henle (2014: 12).

Wichtig ist es jedoch Dehnungen und "lange Konsonanten", wie sie auch in der Kontaktsprache Ungarisch vorkommen, in Gestalt von nachgestellten Doppelpunkten zu notieren (zum Beispiel so:, gekann:t, nit: etc.). Dies betrifft sprechsprachliche Realisierungen, die von der standardsprachbasiert erwarteten Länge abweichen, gilt aber besonders im Falle abweichender Aussprache von deutschen Vokabeln insbesondere bei speziellen Dialektwörtern. Denn durch die Existenz von langen Konsonanten im Ungarischen ist es möglich, dass Sprachkontaktphänomene auftreten, bei denen deutsche (Dialekt-)wörter mit z.T. ungarischer bzw. ungarisch beeinflusster Artikulation, in diesem Falle Dehnung, vorkommen.

Solche Mehrsprachigkeits- oder Kontaktphänomene sind außerdem kursiv darzustellen (z.B. er hatte dann so n *izé*¹³ so n neuen volkswagen). Wann immer notwendig, ist in einer neuen Zeile möglichst direkt unter der entsprechenden Äußerung oder Passage eine Übersetzung in schrägen Klammern vorzunehmen. Wie auch beim simultanen Sprechen lässt sich die Zeile mit Leerzeichen auffüllen (dies wird im nachfolgenden Beispiel veranschaulicht).

Was Überlappungen und simultanes Sprechen betrifft, wird auf eine Präzisierung der Segmentierung durch eckige Klammern verzichtet, d.h., es genügt, wie oben beschrieben, turns bzw. Intonationsphrasen und das Einsetzen der Äußerung in neuen Zeilen und an

¹³ Ungarisches Substantiv, etwa in der Bedeutung von 'Dings', 'Dingsda'.

entsprechender Stelle zu verzeichnen, vgl. auch Hagemann & Henle (2014: 12) und Selting, Auer & Barth-Weingarten u.a. (2009: 364f.).

Beispiel einer Transkription nach den Projektrichtlinien:

I = Interviewerin, weiblich, geboren 1980 in Fünfkirchen/Pécs

B = Befragter, männlich, geboren 1933 in Band, höchster Bildungsgrad Grundschule (4 Klassen), (früherer) Beruf Bauer

Aufnahmeort = Band

Datum = 15. Februar 2016

B: wir hatten dann ehm *megspóroltuk az összes pénzüket*
 /wir hatten all unser geld gespart/
 das war en *nagyobb summa* und ehm *tetszik tudni*
 /ein größerer betrag/ /wissen sie/

I: und davon haben sie

B: ja davon wir kauften unser erstes eigenes haus wie
a filmekben mit amerikai konyha és terasz
 /im film/ /offener küche und terrasse/

I: aber das besitzen sie heute nicht mehr

B: ja: das haben wir später (verkauft) an mein onkel u:nd
 sind nie wieder hin gegangen ja:

der hatte so viele hunde vor denen hatten wir alle angst

I: achso ach
 aber die familie war auch sehr groß richtig

B: na: nein nicht sehr groß

das war nur

also wir sind nur mit mein onkel und tante gekommen

die hatten auch keine kinder

3.6 Das erschlossene sprechsprachliche Material wird, wie erwähnt, in einer digitalen Datenbank – neben verschiedenen Metada-

ten und sonstigen Informationen (wie Ortsgeschichte, Dialektgrundlage etc., vgl. 3.7.2) – über ein Online-Informationsangebot öffentlich zugänglich gemacht. Wissenschaftler(innen) haben dann die Möglichkeit, die Datenbank zu durchsuchen und die Datensätze nach verschiedenen Kriterien zu filtern und zu ordnen sowie Tonaufnahmen wiederzugeben. Die Datenbank soll erweiterbar sein und kann über die Dauer des Projekts ständig neu zugehendes Belegmaterial erfassen. Zu den technischen Voraussetzungen gehört dabei die Verwendung eines CMS (Content Management Systems), das idealerweise alle nötigen Funktionen zur Administration sowohl des Portals als auch der Datenbank aufweist. Durch wissenschaftliche Mitarbeiter(innen) und Hilfskräfte werden Portal und Datenbank administrativ verwaltet, modifiziert und erweitert sowie Datensätze angelegt und neues Material analysiert, kategorisiert und eingepflegt. Dies garantiert ein stetiges Wachstum des Portals im gesamten Projektzeitraum (vgl. Földes 2016b). Die Informationsstruktur und der web-technologische Apparat sollen im Folgenden vorgestellt werden:

3.7 Funktionalitäten der UZSK-Datenbank – vorgesehene Struktur- und Funktionskonzept

3.7.1 Nutzergruppen und -rechte

Die Registrierung soll mit E-Mail und selbst gewähltem Passwort funktionieren. Dazu sind Daten wie Vorname, Name, Postleitzahl, Ort, Einrichtung, ferner eine Zustimmung zu den Nutzungsbedingungen und eine Datenschutzerklärung erforderlich. Die Registrierung erfolgt, nachdem der Link in einer automatisch versandten Bestätigungs-E-Mail angeklickt wurde.

Es gibt drei Nutzergruppen: `Systemadministrator` (mit allen Rechten), `Editor` (mit den Rechten, Datensätze anzulegen/zubearbeiten, und zum Dateiuupload) und `Betrachter` (mit Rechten, bestimmte Inhalte anzuzeigen). Nach erfolgreicher Registrierung ist man automatisch der Gruppe `Betrachter` zugeordnet. Ein `Systemadministrator` (im Weiteren kurz: `Admin`) kann einem registrierten Nutzer höhere Rechte zuweisen, indem er ihn einer anderen Gruppe (`Admin` oder `Editor`) zuordnet oder umgekehrt: Rechte entzieht. Ohne eingeloggt zu sein, sieht man sonst unter dem Menüpunkt "Datenbank" nur die Aufforderung, sich zu registrieren oder sich anzumelden ("Bitte melden Sie sich an"). Eingeloggte Nutzer sehen die Datenbank in der tabellarischen `Standardansicht`. `Admins` und

Editoren sehen zusätzlich den Adminbereich mit den Optionen "Eintrag hinzufügen" und "Beschreibung hinzufügen". Im Adminbereich sehen Admins außerdem die Rechtevergabe. Grafik 2 zeigt eine Tabelle der einzelnen Nutzer und deren Rechte, ohne Passwort, + Zeitpunkt Anmeldung, + Zeitpunkt der letzten Benutzeranmeldung (im Weiteren kurz: Login).

| Funktion/Rechte | Admin | Editor | Betrachter |
|--------------------|-------|--------|------------|
| Tabelle anzeigen | + | + | + |
| Suche nutzen | + | + | + |
| Eintrag anzeigen | + | + | + |
| Datei aufrufen | + | + | + |
| Eintrag hinzufügen | + | + | - |
| Eintrag bearbeiten | + | + | - |
| Datei hochladen | + | + | - |
| Rechte vergeben | + | - | - |

Grafik 2: Veranschaulichung der Zugriffsrechte der einzelnen Gruppen

3.7.2 Datensatz anlegen/bearbeiten¹⁴ und ausgeben: Stichworte zur Programmierung

- Möglichkeit der Eingabe verschiedener Metadaten sowie Uploads (Audiodatei: wav oder mp3, Transkript als XML-Datei).
- Eingabemaske mit Bereichen für Eingabe- bzw. Auswahlfelder.
- 1. Metaheader: Ereignis-ID, Aufnahmedatum (MM/JJJJ), Aufnahmeort (Drop-Down-Menü gespeicherter Orte), Bezeichnung (neu, frei wählbar), alte Bezeichnung (bisheriger Dateiname der Audiodatei, wird beim Upload übernommen, kann nicht geändert werden).
- 2. Explorator: Exploratoren-ID, Geschlecht (Drop-Down: m, w), Geburtsjahr (JJJJ), Herkunftsort (wie bei Aufnahmeort).
- 3. Gewährsperson: Sprecher-ID, Geschlecht (D-D: m, w), Geburtsjahr (JJJJ), Herkunftsort (wie bei Aufnahmeort), Bildungsgrad/Beruf (Text, frei wählbar).

¹⁴ Praktisch dasselbe, was in der Grafik 2 "Eintrag hinzufügen" hieß.

- Beim Klick auf die Schaltfläche "+ Sprecher hinzufügen" erscheint ein weiteres "Befragter"-Feld mit denselben Auswahlfeldern (jeder Sprecher wird unter einer eigenen Sprecher-ID gespeichert).
- Beim Anlegen des Datensatzes wird eine einzigartige Ereignis-ID vergeben (die nicht neu vergeben wird, selbst wenn der Datensatz gelöscht wird) nach dem Format: "UZSK_E_0001" (fortlaufend nummeriert).
- Einzeln gespeichert werden der Datensatz (unter seiner ID), die Information zum Sprecher + Sprecher-ID, die Audiodatei + Audio-ID und das Transkript + Transkript-ID.
- Unter "Eintrag hinzufügen" wird also gespeichert: der Datensatz (der die Verbindung der einzelnen Einträge zueinander und ihre Metadaten festhält), die Sprecher unter ihren IDs, die Audiodatei unter ihrer ID, das Transkript unter seiner ID.
- Die Sprecherdaten sind dabei ereignisübergreifend, d.h., ein Sprecher kann in mehreren Audiodateien und damit Ereignissen/Datensätzen vorkommen.
- Beim Anlegen des Datensatzes (und bei weiterer Bearbeitung) werden zusätzlich folgende Metadaten angelegt, die nur für Nutzer mit Adminrechten in der Bearbeitungsansicht der Datensätze sichtbar sind: Datum Erstellung (TT/MM/JJJJ, HH:MM) durch (Loginname des Nutzers), Datum letzte Bearbeitung (TT/MM/JJJJ, HH:MM) durch (Loginname des Nutzers).
- Beim erstmaligen Klick auf "Speichern" wird der Datensatz angelegt, bei jedem weiteren Klick werden die Daten und die Bearbeitungsdaten aktualisiert.
- Möglichkeit, Datensätze per Klick zu verbergen oder zu veröffentlichen: Datensatz wird beim Klick auf "Speichern" angelegt oder aktualisiert, ist aber nur für Admin/Editor sichtbar.
- Beispiel der Ansicht für den Admin unter "Eintrag hinzufügen" bzw. "Eintrag bearbeiten":

M e t a h e a d e r

| | | | |
|---------------|-------------|------------------|------------|
| ID | UZSK_E_0048 | | |
| Aufnahmedatum | 11/2015 | Aufnahmeort | Band/Bánd |
| Bezeichnung | Band_m_001 | alte Bezeichnung | band02.wav |

E x p l o r a t o r

| | |
|-----------------|---------|
| Exploratoren-ID | horvath |
|-----------------|---------|

Herkunftsort Band/Bánd
 Geschlecht m Geburtsjahr 1987

Gewährsperson
 Sprecher-ID UZSK_S_0025
 Herkunftsort Band/Bánd
 Geschlecht m Geburtsjahr 1933
 Bildungsgrad/Beruf 4. Klasse Grundschule

Feld: + Sprecher hinzufügen

Upload

Audio band02.wav
 Audio-ID UZSK_E_0048_A_00048
 Transkript beleg1938m.xml
 Transkript-ID UZSK_E_0048_T_00048

Bearbeitung

Datum Erstellung 20.04.2016, 11:48 durch schmidt
 Datum letzte Bearbeitung 21.04.2016, 13:25 durch roth

Eintrag verbergen

Speichern

Beispiel – Ausgabe in Tabellenform (d.h. Ansicht, die die Nutzer sehen, mit den Optionen, aus denen sie auswählen können):

| # | ID | Bezeichnung | Ort | Datum |
|------|-------------|--------------|-----------|---------|
| 0038 | UZSK_E_0040 | Herend_m_002 | Herend | 02/2016 |
| 0039 | UZSK_E_0041 | Iklad_w_001 | Iklad | 04/2016 |
| 0040 | UZSK_E_0043 | Iklad_m_002 | Iklad | 05/2016 |
| 0041 | UZSK_E_0044 | Moor_w_008 | Moor/Mór | 04/2014 |
| 0042 | UZSK_E_0048 | Band_m_001 | Band/Bánd | 11/2015 |
| 0043 | UZSK_E_0049 | Band_w_002 | Band/Bánd | 03/2016 |

– Das Rautezeichen # bedeutet: Fortlaufend nummeriert (von 0001 an) bei der Ausgabe, jedoch nicht fest an Datensatz gebunden (da Datensätze verborgen oder gelöscht werden können); dient der Anzeige, wie viele Ergebnisse bei der jeweiligen Suche gefunden wurden bzw. wie viele ohne Suchfilter vorhanden sind, Ordnungsreihenfolge richtet sich nach der ID.

- Wenn man die Ausgabe nicht per Suche einschränkt, werden standardmäßig alle vorhandenen Datensätze angezeigt, Admin/Editor sieht zusätzlich verborgene (besonders hervorgehoben).
- Beim Klick auf die Teile des Tabellenkopfes (Rautezeichen, ID, Bezeichnung etc.) wird die Reihenfolge der Darstellung umgeordnet oder umgekehrt wie folgt: Rautezeichen und ID werden in umgekehrter Reihenfolge angezeigt, Datum in chronologischer oder anti-chronologischer (ältestes, jüngstes) Reihenfolge, Bezeichnung und Ort in alphabetischer oder bei erneutem Klick in umgekehrt alphabetischer Reihenfolge.
- Für Nutzer mit Adminrechten erscheinen in der Tabelle auch die verborgenen Einträge in der Reihenfolge der ID, lassen sich aber ausblenden, um die Tabelle in Nutzeransicht zu sehen.
- Beim Klick auf die jeweilige ID, also nachdem man sich etwas ausgewählt hat, soll der Datensatz nach folgendem Schema erscheinen; Datenansicht:

ID: UZSK_E_0048 Bezeichnung: Band_m_001
 Ort: Band/Bánd Datum: 1/2015

Gewährsperson

Herkunftsort Band/Bánd

Geschlecht m Geburtsjahr 1933

Bildungsgrad/Beruf 4. Klasse Grundschule

Audiodatei (mit Playsymbol › zum Abspielen, das sich dann in ein Pausensymbol || zum Unterbrechen verwandelt, und ein Fortschrittsbalken, der die Position und Zeit sowie die Gesamtdauer anzeigt, so wie bei Programmen für Audio/Video üblich).

Transkript

(Transkripttext falls vorhanden, sonst steht hier "nicht vorhanden")

Dialektbeschreibung

(falls vorhanden, sonst steht hier "nicht vorhanden")

Ortsbeschreibung

(falls vorhanden, sonst steht hier "nicht vorhanden")

Karte

(eingebettete google map).

- Nutzer mit Adminrechten haben zusätzlich zu dieser Darstellung noch eine Schaltfläche "Bearbeiten", um zur Bearbeitungsansicht zu kommen (wie bei "Eintrag hinzufügen").
- Die Anzeigen des Transkripts, der Orts- und Dialektbeschreibung und der Karte können mit einem Klick ein- oder ausgeblendet werden (evtl. ein Symbol daneben und der Text "ausblenden", der sich nach einem Klick in "einblenden" verwandelt).

3.7.3 Suchfunktion (Schaltfläche immer ganz oben sichtbar): in Stichworten

- Auswahlfelder (Drop-Down-Menüs mit allen Möglichkeiten oder "keine Auswahl", falls das Feld nicht in die Suche einbezogen wird): Aufnahmeort, Aufnahmejahr, Geschlecht der Gewährsperson, Herkunftsort der Gewährsperson und Transkript (wenn ein solches vorhanden ist)
- sowie Eingabefeld für Freitextsuche, welches die Transkripte und Metadaten nach Wörtern durchsucht und alle Datensätze anzeigt, in denen das Wort vorkommt.

3.7.4 Ortsansicht (nur Admin und Editor): in Stichworten

- Tabellarische Ansicht der vorhandenen Orte.
- Beim Klick auf einen Ort besteht die Möglichkeit, eine Orts- und eine Dialektbeschreibung ins Textfeld einzutragen (Text, eventuell als html).
- Die obigen Elemente können dann in der Datensatzansicht ausgegeben werden, da beim Anlegen eines Datensatzes ein Ort ausgewählt wird (dieselben Orts- und Dialektbeschreibungen werden in mehreren Datensätzen verwendet).
- Möglichkeit, neue Orte hinzuzufügen: deutscher Name/ungarischer Name, geographische Daten, Ortsbeschreibung (Text), Dialektbeschreibung (Text).

Koordinaten der Orte (für Drop-Down-Menü der Orte und Darstellung in google maps):

| | |
|-----------|---------------------------|
| Band/Bánd | 47°07'22.6"N 17°46'45.6"E |
| Herend | 47°08'02.0"N 17°45'10.4"E |

| | |
|----------------------|---------------------------|
| Iklad | 47°39'45.8"N 19°26'33.2"E |
| Marka/Márkó | 47°07'26.4"N 17°49'02.8"E |
| Moor/Mór | 47°22'29.5"N 18°12'18.8"E |
| Nadasch/Mecseknádasd | 46°13'30.4"N 18°27'51.1"E |
| Putawam/Pusztavám | 47°25'46.1"N 18°13'58.9"E |
| Tiedisch/Töttös | 45°54'51.5"N 18°32'19.3"E |
| Wieland/Villány | 45°52'08.4"N 18°27'10.1"E |

In einer späteren Projektphase werden die Transkripttexte zusätzlich annotiert bzw. aligniert, so dass beim Klick auf verlinkte Wörter die Audiodatei an die entsprechende Stelle springt. Es sollte frühzeitig schon so geplant werden, dass dies später einmal möglich ist.

4 Schluss

Im Projekt geht es insgesamt um die Erschließung sprachkommunikativer Quellen,¹⁵ verbunden mit deren wissenschaftlicher Analyse. Erbracht werden sollen relevante Forschungsleistungen zur Variations- und Kontaktlinguistik. Es soll ein umfassendes funktionsfähiges Datenkorpus erstellt und dieses in einem zeitgemäßen webbasierten System präsentiert werden; gehört doch Sprachdokumentation, die neben Sprachbeschreibung einen der beiden Hauptbereiche der Deskriptiven Linguistik¹⁶ verkörpert, eigentlich zu den Kernaufgaben der Sprachwissenschaft.

Der jeweilige Entwicklungsstand bzw. die Progression ist auf der Projekthomepage unter den folgenden Web-Adressen kontinuierlich zu verfolgen:

www.ungarndeutsch.de,
www.ungarndeutsch.eu,
www.donauschwaebisch.de,
www.donauschwaebisch.eu,
www.uzsk.de und
www.uzsk.eu.

¹⁵ Dabei wird nicht nur eine Sprach-, sondern u.U. auch eine Art Kulturdokumentation entstehen.

¹⁶ Den Konnex zwischen Dokumentation und Beschreibung von Sprachen haben z.B. Himmelmann (1998) und Lehmann (2004) systematisch herausgearbeitet.

Ein weiterer qualitativer und quantitativer Ausbau soll in einer angestrebten Verlängerungsphase des Vorhabens erfolgen.

Literatur

- Bowern, C. (2015): *Linguistic fieldwork. A practical guide*. 2. d. New York [u.a.]: Palgrave Macmillian.
- Chelliah, Sh. (2013): Fieldwork for language description. In: Podesva, R.J. & Sharma, D. (eds.): *Research Methods in Linguistics*. Cambridge: Cambridge University Press, 51-73.
- Comrie, B. & Golluscio, L. (eds.) (2015): *Language Contact and Documentation. Contacto lingüístico y documentación*. Berlin, München & Boston: de Gruyter.
- Dresing, Th. & Pehl, Th. (2015): *Praxisbuch Interview, Transkription & Analyse. Anleitungen und Regelsysteme für qualitativ Forschende*. 6. Aufl. Marburg: Dresing & Pehl.
- Erb, M. (2010): Sprachgebrauch der Ungarndeutschen. Geschichte – Tendenzen – Perspektiven. In: Kostrzewa, F. & V. Rada, R. unter Mitarbeit von Knipf-Komlósi, E. (Hrsg.): *Deutsch als Fremd- und Minderheitensprache in Ungarn. Historische Entwicklung, aktuelle Tendenzen und Zukunftsperspektiven*. Baltmannsweiler: Schneider Hohengehren, 118-146.
- Földes, Cs. (2013): Sprachliche Praktiken im Spannungsfeld von Variation und Mehrsprachigkeit: Ein Beitrag zur Empirie. In: Schneider-Wiejowski, K., Kellermeier-Rehbein, B. & Haselhuber, J. (Hrsg.): *Vielfalt, Variation und Stellung der deutschen Sprache*. Berlin & Boston: de Gruyter, 119-142.
- Földes, Cs. (2016a): Ungarndeutsches Zweisprachigkeits- und Sprachkontaktkorpus: Konzept, Design und Inhalte. *Zeitschrift für interkulturelle Germanistik* 7, 142-157.
- Földes, Cs. (2016b): Diskurse im Wirkungsraum von Zweisprachigkeit. Werkstattbericht aus einem Forschungs- und Dokumentationsprojekt. In: Duś, M., Kołodziej, R. & Rojek, T. (Hrsg.): *Wort – Text – Diskurs*. Frankfurt a.M. [etc.]: Lang, 321-336.

- Hagemann, J. & Henle, J. (2014): *Transkribieren nach GAT 2 (Minimal- und Basistranskript) – Schritt für Schritt*. Freiburg: PH. Im Internet unter https://www.ph-freiburg.de/fileadmin/dateien/mitarbeiter/hagemannfr/Transkribieren_nach_GAT_2.pdf (Zugriff: 25.04.2016).
- Himmelman, N.P. (1998): Documentary and descriptive linguistics. *Linguistics* 36, 161-195.
- Knipf-Komlósi, E. (2011): *Wandel im Wortschatz der Minderheitensprache: am Beispiel des Deutschen in Ungarn*. Stuttgart: Steiner.
- Lasatowicz, M.K. & Pelka, D. (2011): *Sprachbiographien in Oberschlesien*. Berlin: Trafo.
- Lehmann, Ch. (2004): Documentation of grammar. In: Sakiyama, O., Endo, F., Watanabe, H. & Sasama, F. (eds.): *Lectures on endangered languages: 4. From Kyoto Conference 2001*. Osaka: Osaka Gakuin University, 61-74.
- Manherz, K. (2014): A magyarországi német dialektológiáról. Kutatási tendenciák és eredmények. *Magyar Nyelv* 110, 1-16.
- McEnery, T. & Hardie, A. (2012): *Corpus Linguistics. Method, Theory and Practice*. Cambridge: Cambridge Univ. Press.
- Mello, H. (2014): Methodological issues for spontaneous speech corpora compilation. The case of C-ORAL-BRASIL. In: Raso, T. & Mello, H. (eds.): *Spoken Corpora and Linguistic Studies*. Amsterdam & Philadelphia: Benjamins, 27-68.
- Mollay, K. (1982): *Német-magyar nyelvi érintkezések a XVI. század végéig* [Deutsch-ungarische Sprachkontakte bis Ende des 16. Jahrhunderts]. Budapest: Akadémiai Kiadó.
- Muysken, P., Hammarström, H., Birchall, J., van Gijn, R., Krasnoukhova, O. & Müller, N. (2015): Linguistic areas, bottom-up or top-down? The case of the Guaporé-Mamoré. In: Comrie, B. & Golusció, L. (eds.): *Language Contact and Documentation. Contacto lingüístico y documentación*. Berlin, München & Boston: de Gruyter, 205-229.
- Pilarský, J. (2013): *Deutsch-ungarische kontrastive Grammatik*. Debrecen: Debreceni Egyetemi Kiadó.

- Ruoff, A. (1973): *Grundlagen und Methoden der Untersuchung gesprochener Sprache. Einführung in die Reihe "Idiomata" mit einem Katalog der ausgewerteten Tonbandaufnahmen*. Tübingen: Niemeyer.
- Scherer, C. (2014): *Korpuslinguistik*. Zweite, aktual. Aufl. Heidelberg: Winter.
- Schilling, N. (2013): *Sociolinguistic fieldwork*. Cambridge [etc.]: Univ. Press.
- Selting, M., Auer, P. & Barth-Weingarten, D. u.a. (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion* (www.gesprachsforschung-ozs.de). Ausgabe 10, 353-402.
- Tagliamonte, S.A. (2006): *Analysing Sociolinguistic Variation*. 3. printing. Cambridge [etc.]: Univ. Press.

Prof. Dr. Dr. h.c. Csaba Földes
Universität Erfurt, Philosophische Fakultät
Lehrstuhl für Germanistische Sprachwissenschaft
Nordhäuser Straße 63
D-99089 Erfurt
foeldes@foeldes.eu
www.foeldes.eu