

FREDDIE (FREiburg English Dialect Database for Instruction and E-learning)

– Eine multimediale Plattform für forschendes, korpusbasiertes Lernen –

Antragsteller: Prof. Dr. Dr. h.c. Bernd Kortmann, Katja Roller, Marten Juskan

1. Einleitung und Kurzzusammenfassung des Projekts

Das Projekt zielt darauf ab, das linguistische Korpus FRED (*FREiburg English Dialect corpus*) zur multimedialen Lehr- und Lernplattform FREDDIE weiterzuentwickeln. Im Rahmen der IDA-Förderung sollen hierbei drei zentrale Ziele umgesetzt werden: (1) Die FRED-Interviews werden durch Anonymisierung und Alignierung für die Online-Bereitstellung nutzbar gemacht. (2) Es wird ein FREDDIE-Onlineportal erstellt, das Lehr- und Lernmaterialien sowie Analyseprogramme bereitstellt. (3) Die FREDDIE-Materialien werden in der Lehre eingesetzt, erprobt und optimiert. Insgesamt stellt FREDDIE eine nachhaltige und zugleich stets aktualisierbare Ressource dar, mit der Studierende auf vielfältige Weise – in Kursen und im Selbststudium – an das forschende, korpusbasierte Lernen herangeführt werden können.

In ein bis zwei vollständig neu konzipierten Seminaren (und vielen weiteren bestehenden, die von FREDDIE-basierten Modulen profitieren können) werden Studierende an die komplexen Forschungsmechanismen herangeführt, die die aktuelle empirische Linguistik prägen. Sie werden in einem mehrstufigen Verfahren authentische Daten aufbereiten, kritisch hinterfragen, analysieren und visualisieren. Durch schrittweises, handlungsorientiertes Üben werden sich die Studierenden auf diese Weise ein komplexes Portfolio aus fachlich-linguistischen, technischen und Metakompetenzen aneignen, und zwar mit einer Intensität, die bisher – ohne die Ressource FREDDIE – nicht zu realisieren ist.

2. Ausgangssituation

Den Ausgangspunkt für FREDDIE bietet das FRED-Korpus – eine Datenbank mit Aufnahmen und Transkripten im Umfang von 2,5 Millionen Wörtern von Dialektsprechern aus verschiedenen Regionen in England, Schottland und Wales. Das Korpus, das größte seiner Art weltweit, wurde im Rahmen von zwei DFG-Förderungen unter der Leitung von Prof. Bernd Kortmann von 2000 bis 2005 erstellt und wird aktuell in Zusammenarbeit mit der Freiburger Universitätsbibliothek online verfügbar gemacht. Parallel zur Bereitstellung dieser Datenbank für die internationale Forschung kann FRED Online mit IDA-Mitteln auch umfangreich als multimediale, interaktive Plattform für die Lehre aufbereitet werden. Das Korpus bietet vielfältige Möglichkeiten für forschendes Lernen ab dem ersten Semester (in sogenannten *Doing-Linguistics*-Kursen, in denen schon gemäß der gültigen Studienordnungen am Englischen Seminar ein erstes Forschungsprojekt durchgeführt werden muss), in Pro- und Hauptseminaren, der linguistischen Masterausbildung (u.a. im obligatorischen Modul Forschungsdesign) bis hin zur Doktorandenausbildung (etwa für quantitative Analysen in Statistik-Seminaren).

Die geplante Online-Plattform FREDDIE verzahnt bereits bestehende Dienste, etwa bereitgestellt vom Englischen Seminar und der Universitätsbibliothek, mit neuen Anwendungsmöglichkeiten, Konzeptionen und Programmen, die mit Hilfe von IDA-Mitteln erstellt werden. Zu den bestehenden Diensten zählt beispielsweise das FRED-S-Korpus, ein Teilkorpus von FRED im Umfang von etwa 1 Million Wörtern, das bereits seit kurzem über FreiDok online zugänglich ist (<https://www.freidok.uni-freiburg.de/data/10208>). Es ist also schon möglich, auf Transkripte von Interviews aus verschiedenen Regionen der Britischen Inseln zuzugreifen und die Texte im Hinblick auf bestimmte Wörter zu durchsuchen. Hierfür kann die interaktive Datenbank herangezogen werden (<https://fred.ub.uni-freiburg.de/>), oder, für tiefergehende Analysen, die Emulation des Programms AntConc (<https://www.freidok.uni-freiburg.de/data/10845>), welche in Zusammenarbeit der Universitätsbibliothek, des Englischen Seminars und des Rechenzentrums erstellt wurde. Bisher ist die Nutzung von FRED(-S)

allerdings auf die Analyse von Texten beschränkt, da (1) ein Großteil der Aufnahmen noch nicht anonymisiert ist (und daher aus Datenschutzgründen nicht online bereitgestellt werden kann) und (2) noch keine Alignierung vorgenommen wurde, Texte und Aufnahmen also unverbunden nebeneinander stehen. Zudem bietet FRED zwar theoretisch viele Möglichkeiten für die Lehre, eine systematische Entwicklung und Bereitstellung von Lehrmaterialien und –konzepten wurde allerdings bislang nicht vorgenommen.

3. Ziele im Rahmen der IDA-Förderung

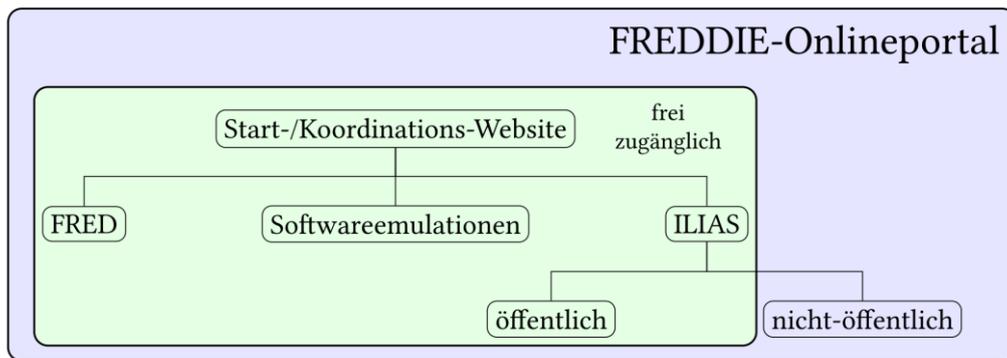
Mithilfe der IDA-Förderung sollen drei zentrale Ziele im Rahmen von FREDDIE verfolgt werden: (1) Die Anonymisierung und Alignierung der FRED-Interviews, (2) der Aufbau eines FREDDIE-Onlineportals, das Lehr- und Lernmaterialien sowie Analyseprogramme bereitstellt, sowie (3) die Implementierung von FREDDIE-Materialien in der Lehre.

Im Hinblick auf (1) können sowohl die Anonymisierung also auch die Alignierung von Texten und Aufnahmen mithilfe des kostenlosen Programms PRAAT erledigt werden, sind allerdings sehr zeitintensiv. Um dies zu realisieren, möchten wir daher gerne eine IDA-finanzierte geprüfte Hilfskraft einstellen. Die Online-Bereitstellung der anonymisierten Aufnahmen und die Alignierung werden die Einsatzmöglichkeiten von FRED in der Lehre massiv ausweiten. Durch die Alignierung der Daten können die Tonaufnahmen beispielsweise mit zeitlich passenden Untertiteln versehen werden. Dies ist nicht nur besonders hilfreich bei Aufnahmen mäßiger Tonqualität und undeutlich artikulierender Sprecher, sondern ermöglicht auch das direkte Springen zu einem bestimmten Abschnitt der Aufnahme (realisiert über eine interaktive Liste der Untertitel – dies kann mit der Software Camtasia umgesetzt werden). Schon diese beiden Punkte werden den Wert von FRED für die Lehre spürbar erhöhen, da es mit dem Alignment deutlich einfacher und komfortabler wird, in einer Lehrveranstaltung „echte“ Daten auszugsweise zu Illustrationszwecken vorzuführen.

Durch den (2) Aufbau und die (3) Implementierung des FREDDIE-Onlineportals wird der Nutzen für die Lehre allerdings noch einmal kräftig gesteigert. FREDDIE muss hierbei als eine Klammer verstanden werden, die Bestehendes und neu zu Schaffendes verbindet und zueinander in Bezug setzt. Nach Fertigstellung wird FREDDIE aus folgenden Einzelmodulen bestehen:

- a) FRED (FreiDok plus), die zugrunde liegende Datenbank mit den Transkripten und alignierten Audiodateien
- b) Emulationen von Shiny (für Statistik- und Visualisierungsapps - separater Server/virtuelle Maschine vom RZ) und AntConc (Korpusanalysen – RZ/UB)
- c) Allgemeines Lehr- und Lernmaterial, Tutorials zum Selbststudium, Analyseskripte etc. (ILIAS – öffentlicher Bereich)
- d) Konkrete und einer bestimmten Lehrveranstaltung zugeordnete Tests, Wikis, etc. (ILIAS – abgeschlossener Bereich)
- e) Eine einfache Website, die als übergeordnete Startseite das Gesamtkonzept erläutert und die einzelnen Bausteine verlinkt (dies wird sein: freddie-dialects.org)

Mit Ausnahme von d) wird sämtliches Material frei übers Internet zugänglich sein (siehe Grafik).



Mit Hilfe der in diesem Rahmen zur Verfügung gestellten Unterrichtsmaterialien und Analyseprogramme können Studierende dann einen Forschungsprozess von Beginn an Schritt für Schritt erleben. Beispielfähig sei hier eine phonetisch/phonologische Analyse genannt, bei der Forscher zunächst Tonaufnahmen transkribieren müssen, um anschließend relevante Messungen akustischer Parameter vorzunehmen und diese Werte schließlich auswerten zu können. Die Simulation in der Lehre – als Beispiel für den „Instruction“-Anteil an FREDDIE – soll nun folgendermaßen aussehen: Jeder Studierende fertigt als ersten Schritt ein grobes orthographisches Transkript eines kurzen Interviewschnipsels (ca. 90 Sekunden) aus FRED an. Diese Minitranskripte werden auf ILIAS (im abgeschlossenen Bereich) eingereicht, mithilfe des integrierten peer-feedback Features innerhalb der Gruppe der Lernenden überarbeitet und die endgültige Version von der/dem Lehrenden nochmals auf zufriedenstellende Qualität überprüft. Der/die Lehrende kann zur Korrektur auf die bereits bestehenden, alignierten Transkripte der FRED-Interviews zurückgreifen. Es soll hierbei also nicht darum gehen, Studierende als unbezahlte Arbeitskräfte am FRED-Korpus zu gewinnen, sondern vielmehr um einen didaktischen Dienst an den Studierenden. Aus den groben Transkripten wird anschließend ein sehr viel feineres Alignment automatisch generiert (diese Funktionalität wird von der University of Pennsylvania kostenlos über das Internet zur Verfügung gestellt). Das Ergebnis dieser automatischen Weiterverarbeitung diskutieren und analysieren die Studierenden kritisch (zum Beispiel mittels eines ILIAS-Wikis) und nehmen gegebenenfalls manuelle Korrekturen vor. Für jeden Interviewausschnitt, der diesen Prozess durchlaufen hat, gibt die/der Lehrende schließlich das ausführliche Alignment des gesamten zugehörigen Interviews frei, das vorher am Lehrstuhl (mit IDA-Mitteln) erstellt worden ist. Nun können Studierende unter Zuhilfenahme von auf dem FREDDIE-Onlineportal zur Verfügung gestellter Software(skripte) relevante Messungen automatisiert vornehmen und die Ergebnisse auch gleich visualisieren. Auf der Basis dieses Materials wird schließlich eine schriftliche Analyse verfasst.

Der Einsatz von FREDDIE wird also die Studierenden sehr konkret und zielgerichtet auf die Verfassung einer empirischen wissenschaftlichen Hausarbeit hinführen. Vor allem aber werden die Lernenden auf diesem Weg sämtliche Phasen durchlaufen, die auch bei einer „echten“ Forschungsarbeit vonnöten sind (mit Ausnahme des eigentlichen Datensammelns – im Beispiel das Führen und Aufzeichnen der Interviews). Normalerweise würde ein solches Vorgehen bedeuten, dass die Studierenden aus Zeitgründen nur mit einer sehr kleinen, eher anekdotischen Datensammlung arbeiten müssten. Mit FREDDIE werden die meisten Schritte hinsichtlich des quantitativen Aspekts jedoch nur angedeutet und die Studierenden für qualitativ saubere Arbeit im Kleinen mit einem Vielfachen der tatsächlich selbst erhobenen Daten „belohnt“. In der Folge können sie dann mit einem vollwertigen Datensatz weiterarbeiten, wodurch viele konkrete Fragestellungen überhaupt erst untersuchbar werden. Durch die aktive und kritische Auseinandersetzung mit authentischen Forschungsdaten können Kompetenzen in den Bereichen problemorientiertes, projekt-, erfahrungs- und praxisbezogenes Lernen gefördert werden. Zudem honoriert das geplante Verfahren kooperatives Vorgehen: Koordination und Arbeitsteilung bedeuten im Ergebnis mehr Daten für die gesamte Gruppe – wie im tatsächlichen Forschungsalltag.

Zusätzlich zu den Lehrmaterialien ist umfangreiches Dokumentationsmaterial in Text- und Videoform geplant – als ein Teil der „E-Learning“-Komponente von FREDDIE. Dadurch soll ermöglicht werden, dass sich Studierende Fähigkeiten zur Nutzung der Funktionen und Möglichkeiten von FREDDIE (auch) im autonom organisierten Selbststudium aneignen. Zu diesen Funktionen gehören unter anderem die bereits verfügbare Emulation der Software AntConc (zur Durchführung von Korpusanalysen direkt auf der Website, <https://www.freidok.uni-freiburg.de/data/10845>) sowie die (geplante) Einbindung des Programms Shiny. Shiny ist eine frei verfügbare Software, mit deren Hilfe unkompliziert und ohne nennenswerte Vorkenntnisse das Potential der (ebenfalls frei verfügbaren) Statistiksoftware R genutzt werden kann. Auf diese Weise werden Studierende in der Lage sein, ihre Daten auf der FREDDIE-Website hochzuladen und anschließend mit wenigen Klicks einfache statistische Tests durchzuführen und passende Schaubilder zu erstellen, die dann wiederum direkt exportiert und in die eigene Hausarbeit eingefügt werden können. So soll ein weiterer Aspekt (quantitativer) empirischer Forschung für die Lehre zugänglicher gemacht werden.

4. Innovationscharakter, Nachhaltigkeit, Transferpotential

Das innovative FREDDIE-Konzept ermöglicht es, linguistische Forschung mithilfe von authentischen Daten und modernsten statistischen Analyseprogrammen sowie Visualisierungsmethoden, gebündelt in einem umfassenden E-Learning-Portal, erfahrbar zu machen. Das gesamte Projekt ist in höchstem Maße auf Nachhaltigkeit ausgelegt, da im Zentrum die Erstellung einer Datenbank mit umfangreichen Funktionen und Anwendungsmöglichkeiten steht. Die speziell auf FREDDIE zugeschnittenen Lehrveranstaltungen werden ihren Fokus auf der zu erlernenden Methode (und weniger auf einem Sachthema) haben, sind deshalb weitgehend unabhängig von der ausführenden Lehrperson und werden über Jahre (mit nur minimalen Anpassungen) aktuell bleiben. Vor allem aber wird alles, was im Laufe des Projektes erstellt wird – das Portal selbst, die Lehrmaterialien, Tutorial-Videos und Anwendungsmöglichkeiten – auch nach Ende der Finanzierung bereitstehen und so weltweit genutzt werden können. Die Verankerung von FREDDIE in der Freiburger Lehre wird zudem durch Curricula und die diversen gültigen (tw. gerade frisch beschlossenen bzw. aktualisierten) Prüfungsordnungen unterstützt, beispielsweise im Rahmen der *Doing-Linguistics*-Kurse im Bachelorstudiengang *English and American Studies* oder der Veranstaltungen zu theoretischen und praktischen Grundlagen linguistischer Forschung für Masterstudierende (in den Studiengängen (*European*) *Linguistics* und *English Language and Linguistics*). Durch die Kooperation mit der Universitätsbibliothek und dem Rechenzentrum (beispielsweise bei der bereits bestehenden AntConc-Emulation) wird die interdisziplinäre Zusammenarbeit gefördert. Die oben skizzierte Aufbereitung der FRED-Interviews bietet zudem nicht nur in der linguistischen Lehre ganz neue Möglichkeiten. FREDDIE kann beispielsweise auch in der Geschichtswissenschaft (etwa in der Sozialgeschichte) und Literaturwissenschaft (Erzählforschung, Narratologie) genutzt und für deren spezielle Zwecke adaptiert werden. Auch der Einsatz in der gymnasialen Oberstufe ist denkbar. Generell bietet die Online-Bereitstellung des FREDDIE-Portals die Möglichkeit, die Materialien und Programme an vielen (Hoch-) Schulen im In- und Ausland zu nutzen.

Im Rahmen des multimedialen FRED-Ansatzes können Studierende weltweit auf international bislang einmalige Weise zum forschenden Lernen (zunächst angeleitet, schnell aber auch autonom) befähigt und motiviert werden und insgesamt Kompetenzen in einem Kernbereich der Digital Humanities erwerben.