

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (<http://www.ub.tuwien.ac.at>).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (<http://www.ub.tuwien.ac.at/englweb/>).



FAKULTÄT  
FÜR INFORMATIK

Faculty of Informatics

# GESTU

**Evaluierung von technischen Hilfsmitteln zur Förderung Studierender mit Hörbehinderung im österreichischen tertiären Bildungssektor und Einführung geeigneter Technologien an der TU Wien**

## DIPLOMARBEIT

zur Erlangung des akademischen Grades

### Diplom-Ingenieur

im Rahmen des Studiums

### Software Engineering & Internet Computing

eingereicht von

**Werner Hans Josef Nemecek, BSc**

Matrikelnummer 0425637

an der

Fakultät für Informatik der Technischen Universität Wien

„Zentrum für Angewandte Assistierende Technologien“ (AAT) vom Arbeitsbereich

„Human Computer Interaction“ (HCI) des Instituts für „Gestaltungs- und Wirkungsforschung“

Betreuung

Betreuer: Ao.Univ.-Prof. Dipl.-Ing. Dr.techn. Wolfgang Zagler

Mitwirkung: Dipl.-Ing. Georg Edelmayer

Wien, 18.03.2013

\_\_\_\_\_  
(Unterschrift Verfasser)

\_\_\_\_\_  
(Unterschrift Betreuer)



# Erklärung zur Verfassung der Arbeit

Werner Hans Josef Nemecek, BSc  
Hafengasse 10/11, 1030 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

---

(Ort, Datum)

---

(Unterschrift Verfasser)



# Danksagung

An dieser Stelle bedanke ich mich besonders bei meinen Eltern, Großeltern, Tante und Onkel Fiala und meiner gesamten Familie. Sie haben mich während meines gesamten Bildungsweges stets gefördert und mir durch mein Studium eine interessante und sehr lehrreiche Zeit ermöglicht. Ohne diesen Lebensabschnitt und die prägenden Erfahrungen wäre ich heute sicher nicht der selbe Mensch. Dafür möchte ich mich bei ihnen allen für ihre Unterstützung aber auch Geduld bedanken.

Meiner Verlobten danke ich für ihren Beistand, guten Zuspruch und besonders ihr Verständnis während der, für mich schwierigen und anstrengenden Zeit des Schreibens der Diplomarbeit.

Großer Dank gebührt meinem Betreuer, Prof. Dr. Wolfgang Zagler. Er hat mir dieses spannende Thema anvertraut und mir dadurch die Mitarbeit am GESTU Projekt ermöglicht. Er unterstützte mich bei der Diplomarbeit stets über die Maßen mit seinen Anregungen, Ideen und Antworten. Ebenso bedanke ich mich bei Dipl.-Ing. Georg Edelmayer, mit dem ich die Gelegenheit hatte, während des Modellversuchs zusammenzuarbeiten. Er half mir bei der Umsetzung des Themas meiner Diplomarbeit durch sein Wissen um wissenschaftliches Arbeiten.

Weiters bedanke ich mich bei allen Teilnehmern/-innen des GESTU Projekts und dem gesamten Team für die gute Zusammenarbeit und die interessante und lehrreiche Zeit.

Mein Dank gilt auch meinen Freunden, speziell Christine Braunsteiner für den Austausch über das Schreiben von Abschlussarbeiten und die Hilfe die sie mir dadurch bot. Ich bedanke mich auch bei Allen, mit denen ich während der Zeit meiner Diplomarbeit zahlreiche anregende und bereichernde Gespräche rund um die Themen Studium, Abschlussarbeiten und Arbeit geführt habe.

Großen Dank spreche ich einem besonderen Studienkollegen - Christian Hattinger - aus. Er hat mich vor nicht ganz drei Jahren mit dem spannenden Themenbereich in Berührung gebracht. Unsere Zusammenarbeit und der rege Austausch während unserer Mitarbeit am GESTU Projekt sowie beim Schreiben meiner Diplomarbeit war für mich sehr hilfreich, wichtig und bereichernd.



# Kurzfassung

Der von Juli 2010 bis Juli 2012 durchgeführte Modellversuch GESTU (Gehörlos erfolgreich studieren an der TU Wien) hatte es zum Ziel bestehende Hürden für gehörlose sowie schwerhörige Studierende abzubauen. Das sollte nicht nur mit klassischen Mitteln ermöglicht werden, sondern auch unter Zuhilfenahme technischer Hilfsmittel. Dazu wurden zu Beginn des Projekts die gängigen Methoden zur Unterstützung hörbehinderter Studenten/-innen erfasst und in der Diplomarbeit von Hattinger [Hat11] dokumentiert. Ausgehend von seinen Empfehlungen wurden im Zuge der Arbeit die vielversprechendsten Methoden ausgewählt und im laufenden Lehrveranstaltungsbetrieb an den Wiener Universitäten erprobt und evaluiert.

Das am häufigsten eingesetzte technische Hilfsmittel war die Aufzeichnung der Gebärdensprachdolmetscher/-innen in den Lehrveranstaltungen. Bei der Erprobung der verschiedensten Durchführung- und Einsatzmöglichkeiten und besonders bei der sehr positiven Bewertung durch die hörbehinderten Studierenden zeigte sich der oftmals nicht sofort ersichtliche Beitrag zum Abbau von bestehenden sprachlichen Barrieren.

Eine Vorlesung kann jedoch nicht nur durch Simultandolmetschern/-innen sondern auch durch den Einsatz von Untertiteln zugänglich gemacht werden. Es existieren verschiedene Techniken, mit denen sie in Echtzeit erstellt werden können. Bisher wurde noch keine im Alltagsbetrieb im österreichischen tertiären Bildungssektor eingesetzt wurde. Eine der Erzeugungsmethoden ist das sogenannte Respeaking, das auch während des GESTU Projekts mit großem Erfolg erprobt wurde. Es zeigte sich nicht nur die grundsätzliche Einsetzbarkeit in bestimmten Lehrveranstaltungstypen an den Wiener Universitäten sondern auch, dass es damit den hörbehinderten Studierenden möglich war, dem Inhalt der Vorlesungen zu folgen.

Bei den Aufnahmen ist es auch möglich die Untertitel nachträglich hinzuzufügen. In diesem Zusammenhang ist die automatische und kostengünstige Erstellung durch Spracherkennungssoftware interessant. Daher wurde im Zuge der Diplomarbeit versucht, ein sprecherunabhängiges Spracherkennungssystem für den österreichischen Dialekt zu trainieren und somit zur Transkription von Lehrveranstaltungen einsetzbar zu machen.

Der letzte Schwerpunkt der Arbeit war die Konzeptionierung einer Plattform zur Sammlung von Fachgebärden. Durch die Einbindung in die diversen Aktivitäten des GESTU Projekts zur Sammlung von Gebärden konnte schnell ein verhältnismäßig großer Grundstock an Einträgen gesammelt werden. Bei der Verwendung zeigten sich trotz der erfolgreichen Umsetzung einige Schwächen in den gewählten Lösungsansätzen, aber auch bereits mögliche Verbesserungen.

Durch die Erprobung der ausgewählten technischen Hilfsmittel konnten wichtige Erkenntnisse gewonnen und ein Beitrag zum Abbau bestehender Barrieren für hörbehinderte Studierende an den Wiener Universitäten geleistet werden.



# Abstract

The pilot project GESTU („Gehörlos erfolgreich studieren an der TU Wien“, in English „Successful Deaf Studies at the Vienna University of Technology“) lasted from June 2010 until July 2012 and was initiated to enable more access to studies, for hard of hearing and deaf students. Using classical methods such as interpreters using sign-language was one way to achieve this goal. The other, was the usage of technical aids. Hattinger gathered the common aids for hearing impaired students in his thesis [Hat11] at the beginning of the project. Based on his work the most promising technical aids were chosen and tested in courses at universities in Vienna. Afterwards every method was evaluated in cooperation with the hard of hearing and deaf students participating in the GESTU project.

Lecture capture - focusing on the sign language interpreters - was the most used technical aid. During the course of the experiments various applications and realizations of this method were tested. When asked to judge deaf students generally viewed the quality of services as „very good“. Additionally, some were able to point out further, not immediately obvious, contributions that had helped to diminish existing language barriers.

Using subtitles is another way of making a lecture in tertiary education more accessible for hard of hearing or deaf students. There are currently various methods for creating them in real-time, but none of these were used in courses so far. One of these is the so called respoken technique which was tested with great success during the GESTU project. The evaluation showed the possible use of respoken in everyday lectures and allowing hearing impaired students to follow the topics of the courses.

With respoken creating subtitles in real-time is one option. Another is generating them for a lecture recording. Utilizing speech recognitions to accomplish this, in an automatic and cost-efficient way, is very appealing. Therefore, attempts were made to train a piece of software to achieve good results for the Austrian dialect of the German language when creating subtitles for recorded lectures.

The last focus of this thesis was planning and designing an internet platform for gathering, discussing and searching for missing signs for scientific expressions. The integration of the so called „GESTU Fachgebärden Wiki“ into the various activities around missing signs for scientific expressions in the GESTU project allowed the collection of a relatively large amount of entries. Despite the successful realization the use of this platform showed some flaws in the chosen approach, but also, some possible improvements to fix them.

Finally the experiments with the selected technical aids allowed some important insights and contributed to make lectures at the universities in Vienna more accessible for hard of hearing and deaf students.



# Inhaltsverzeichnis

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Einleitung</b>   | <b>1</b>  |
| 1.1      | Gliederung dieser Arbeit . . . . .  | 2         |
| 1.2      | Begriffsdefinitionen Behinderung, schwerhörig und gehörlos . . . . .              | 2         |
| 1.3      | GESTU - gehörlos erfolgreich studieren an der TU Wien . . . . .                   | 3         |
| 1.4      | Motivation für diese Diplomarbeit . . . . .                                       | 4         |
| 1.5      | Abgrenzung dieser Diplomarbeit . . . . .  | 6         |
| <b>2</b> | <b>Untertitel</b>   | <b>7</b>  |
| 2.1      | Geschichtlicher Hintergrund von Untertiteln und deren aktuelle Einsätze . . . . . | 8         |
| 2.2      | Verschiedene Arten von Untertiteln . . . . .                                      | 9         |
| 2.3      | Methoden zur Untertitelerzeugung . . . . .  | 11        |
| 2.4      | Qualität von Untertiteln bestimmen . . . . .                                      | 14        |
| <b>3</b> | <b>Spracherkennungssoftware</b>   | <b>17</b> |
| 3.1      | Stand der Technik und Forschung . . . . .   | 18        |
| 3.1.1    | Unterteilung von Spracherkennungssystemen . . . . .                               | 19        |
| 3.1.2    | Lösungsansätze für Spracherkennungssysteme . . . . .                              | 21        |
| 3.1.3    | Überblick über aktuelle Spracherkennungssoftware . . . . .                        | 23        |
| 3.2      | Training der EML ASR im Zuge des GESTU Projekts . . . . .                         | 24        |
| 3.2.1    | Durchführung . . . . .  | 25        |
| 3.2.2    | Ergebnisse . . . . .  | 26        |
| 3.2.3    | Empfehlungen . . . . .  | 27        |
| <b>4</b> | <b>Respeaking</b>   | <b>29</b> |
| 4.1      | Geschichte des Respeakings . . . . .  | 30        |
| 4.2      | Definition „Respeaking“ . . . . .   | 30        |
| 4.3      | Respeaking im GESTU Projekt . . . . .   | 32        |
| 4.3.1    | Durchführung . . . . .  | 33        |
| 4.3.2    | Erkenntnisse . . . . .  | 37        |
| 4.3.3    | Empfehlungen . . . . .  | 41        |
| <b>5</b> | <b>Lehrveranstaltungsaufzeichnung</b>   | <b>43</b> |
| 5.1      | Stand der Technik und Forschung . . . . .   | 44        |

|          |  |            |
|----------|--|------------|
| 5.1.1    | Motivation für den Einsatz von Vorlesungsaufzeichnungen . . . . .  | 45         |
| 5.1.2    | Nutzungsszenarien für Lehrveranstaltungsaufzeichnungen . . . . .   | 46         |
| 5.1.3    | Bedürfnisse der Anwender/-innen und Nutzer/-innen . . . . .  | 47         |
| 5.1.4    | Arten der Aufzeichnung . . . . .   | 49         |
| 5.1.5    | Wahrnehmung, Nutzungsverhalten und Auswirkung . . . . .  | 51         |
| 5.1.6    | Lehrveranstaltungsaufzeichnungen für hörbehinderte Studierende . . .   | 54         |
| 5.2      | Synote . . . . .   | 57         |
| 5.3      | Lehrveranstaltungsaufzeichnungen im GESTU Projekt . . . . .  | 59         |
| 5.3.1    | Durchführung . . . . .   | 59         |
| 5.3.2    | Erkenntnisse . . . . .   | 66         |
| 5.3.3    | Empfehlungen . . . . .   | 71         |
| <b>6</b> | <b>Fachgebärdenlexika</b>  | <b>73</b>  |
| 6.1      | Stand der Technik und Forschung . . . . .  | 74         |
| 6.1.1    | Problem fehlender Gebärden für lautsprachliche Ausdrücke . . . . .   | 74         |
| 6.1.2    | Herausforderungen bei der Erstellung von Gebärdenlexika . . . . .  | 75         |
| 6.1.3    | Fachgebärdensammlung . . . . .   | 79         |
| 6.1.4    | ASL-STEM Forum . . . . .   | 80         |
| 6.2      | Sammlung von Fachgebärden im Zuge des GESTU Projekts . . . . .   | 84         |
| 6.2.1    | Realisierung der Plattform zur Sammlung von Fachgebärden . . . . .   | 84         |
| 6.2.2    | Erkenntnisse . . . . .   | 89         |
| 6.2.3    | Empfehlungen . . . . .   | 92         |
| <b>7</b> | <b>Remote Gebärdensprachdolmetschen</b>  | <b>93</b>  |
| 7.1      | Stand der Technik und Forschung . . . . .  | 94         |
| 7.2      | Remote Gebärdensprachdolmetschen im GESTU Projekt . . . . .  | 96         |
| 7.2.1    | Durchführung . . . . .   | 97         |
| 7.2.2    | Erkenntnisse . . . . .   | 97         |
| 7.2.3    | Empfehlungen . . . . .   | 99         |
| <b>8</b> | <b>Schlusswort und Ausblick</b>  | <b>101</b> |
| <b>9</b> | <b>Anhang</b>  | <b>105</b> |
| 9.1      | Transkriptvergleiche einer ausgewählten Vorlesung . . . . .  | 106        |
| 9.1.1    | Das von der trainierten EML ASR erstellte Transkript . . . . .   | 106        |
| 9.1.2    | Das mithilfe der Respeaking Technik erstellte Transkript . . . . .   | 109        |
| 9.2      | Leitfäden für die eTutoren/-innen . . . . .  | 112        |
| 9.2.1    | eTutor Leitfaden für die Durchführung von Respeaking . . . . .   | 113        |
| 9.2.2    | eTutor Leitfaden für die Lehrveranstaltungsaufzeichnung . . . . .  | 115        |
| 9.2.3    | eTutor Leitfaden für die Nachbearbeitung und das Onlinestellen von<br>Lehrveranstaltungsaufzeichnungen . . . . . | 121        |
|          | <b>Abbildungsverzeichnis</b>   | <b>130</b> |

*INHALTSVERZEICHNIS*

xi

**Tabellenverzeichnis**

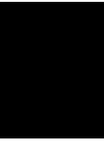
**131**

**Literaturverzeichnis**

**133**



KAPITEL 1



**Einleitung**

## 1.1 Gliederung dieser Arbeit

Im Rahmen dieser Diplomarbeit wurde eine breite Palette verschiedenster technischer Hilfsmittel zur Unterstützung hörbehinderter Studierender erprobt und evaluiert. Aus diesem Grund sind manche der behandelten Themengebiete sehr unterschiedlich. Deshalb werden die einzelnen Maßnahmen in jeweils einem Kapitel behandelt. Diese sind anhand thematischer Zusammenhänge sortiert. In den einzelnen Abschnitten wird zuerst der Stand der Technik und Forschung sowie die benötigten Begriffe und Termini behandelt. Im Anschluss daran wird die Motivation und Durchführung im Zuge des GESTU Projekts sowie die daraus gewonnen Erkenntnisse und Empfehlungen zu den verschiedenen Unterstützungsmethoden beschrieben.

## 1.2 Begriffsdefinitionen Behinderung, schwerhörig und gehörlos

Zu Beginn werden die Begriffe Behinderung, schwerhörig und gehörlos für die weitere Verwendung in der Arbeit kurz definiert. Eine ausführlichere Definition und Behandlung der Begriffe finden sich in der Arbeit von Hattinger (siehe [Hat11, S: 9 - 23]).

Früher wurden Behinderungen allein anhand der medizinischen Fakten bemessen. In den letzten Jahren geht das Hauptaugenmerk bei der Definition weg von der reinen medizinischen Klassifikation und hin zur umfeldorientierten Bewertung. Dabei stellt nicht nur die defizitorientierte Einteilung die alleinige Grundlage zur Beurteilung der Behinderung einer Person dar, sondern auch ob und wie sehr sie dadurch eingeschränkt ist. Die Auswirkungen einer Behinderung treten meistens in einem speziellen Umfeld oder in speziellen Situationen zutage. Das spiegelt sich auch in den neuesten Definitionen der WHO (World Health Organization) in der ICF (International Classification of Functioning, Disability and Health) wieder. Sie kombiniert ein soziales Modell mit den medizinisch und defizitorientierten Fakten (siehe [Hat11, S: 9 - 10]). Das soziale Modell wird folgendermaßen beschrieben:

„Das soziale Modell der Behinderung hingegen betrachtet Behinderung hauptsächlich als ein gesellschaftlich verursachtes Problem und im wesentlichen als eine Frage der vollen Integration Betroffener in die Gesellschaft“ [WHO05].

Das spiegelt auch die Sicht von hörbehinderten Personen auf sich selbst wider. Sie sehen sich meist nur in einer lautsprachlich orientierten Umgebung mit Hürden konfrontiert. Grundsätzlich ist die Festlegung einer Behinderung durch andere Personen problematisch. Das können die Betroffenen nur für sich selbst vornehmen.

Da der Schwerpunkt der Diplomarbeit technische Hilfsmittel für hörbehinderte Personen sind, ist es oftmals erforderlich, die Rahmenbedingungen beziehungsweise die Zielgruppe klar benennen zu können. Daher wird hier genauso wie in den anderen beiden Diplomarbeiten im Zuge des GESTU Projekts folgende Definition von Hattinger verwendet:

„[...] die Begriffe Hörbehinderung sowie Hörbeeinträchtigung für Personen verwendet werden, bei welchen eine Minderung des Hörvermögens besteht. Hörbehinderung und Hörbeeinträchtigung werden weiters als Synonym verwendet und dienen als Überbegriff für schwerhörige sowie gehörlose Menschen. Als gehörlos werden

in dieser Diplomarbeit hörbeeinträchtigte Personen bezeichnet, für welche das bevorzugte Kommunikationsmittel eine Gebärdensprache wie ÖGS (Österreichische Gebärdensprache) ist. Jene Personen, die hörbeeinträchtigt sind und deren bevorzugte Sprache eine Lautsprache ist, werden als schwerhörige Menschen bezeichnet“ [Hat11, S: 1].

### 1.3 GESTU - gehörlos erfolgreich studieren an der TU Wien

An den österreichischen Universitäten studierten im Wintersemester 2011/12 rund 220.000, aber schätzungsweise nur 30 gehörlose Studierende (siehe [Aig12], [Sta]). Schätzungen zufolge leben in Österreich rund 10.000 gehörlose Personen. Das bedeutet - prozentual gesehen - studieren ~2,65% aller Staatsbürger/-innen aber nur ~0,3% aller gehörlosen Personen an öffentlichen Universitäten. Die sehr geringe Anzahl kommt daher, dass die Zulassung zum tertiären Bildungsbereich für hörbehinderte Personen ein sehr großes Hindernis darstellt. Nach der Zulassung zum Studium sind sie mit zum Teil noch größeren Hürden konfrontiert. Die offensichtliche sprachliche Barrieren in einem lautsprachlich orientierten Bildungssystem stellt zwar die größte Barriere dar, ist jedoch bei weitem nicht die einzige. Als Folge dessen benötigen sie für den Abschluss eines Studiums im österreichischen tertiären Bildungsbereich wesentlich mehr Aufwand und vor allem Zeit.

Um die Situation der hörbehinderten Studierenden zu verbessern wurde im Juli 2010 der bis Juli 2012 dauernde Modellversuch GESTU<sup>1</sup> (gehörlos erfolgreich studieren an der TU Wien) gestartet. Er hatte das Ziel, die bestehenden Hürden abzubauen und somit einen besseren Studienerfolg sowie eine kürzere Studiendauer zu ermöglichen. Auf lange Sicht soll somit die niedrige Akademikerquote<sup>2</sup> unter den gehörlosen Österreichern/-innen gesteigert werden.

Das Projekt hatte drei Schwerpunkte - die pädagogische, organisatorische und technische Unterstützung (siehe [Hat11, S: 4 - 5]). Ausgehend davon wurde zu Beginn des Modellversuchs die sogenannte GESTU Servicestelle etabliert. Dort erhalten alle hörbehinderten Studierenden in allen Belangen rund um das Studium Beratung. Insgesamt nahmen 15 hörbehinderte Studenten/-innen am Projekt teil. Sie erhielten eine intensive Betreuung und Unterstützung. So wurden allen Teilnehmern/-innen bis zu fünf Lehrveranstaltungen pro Semester zugänglich gemacht. Weiters wurden die dafür benötigten Gebärdensprachdolmetscher/-innen durch die Servicestelle organisiert. Somit konnten sich die am Modellversuch teilnehmenden Studierenden besser auf ihr Studium konzentrieren. Im Vergleich dazu erhalten hörbehinderte Studenten/-innen normalerweise für lediglich ein bis zwei Lehrveranstaltungen pro Semester finanzielle Unterstützung zur Bezahlung der ÖGS Dolmetscher/-innen (siehe [KS07, S: 352]) und müssen dies auch noch selbst organisieren.

Weiters wurden die Studierenden auch durch Tutoren/-innen, vorrangig Mitschreibtutoren/-innen unterstützt. Das ist notwendig, da es sich bei Gebärdensprachen um visuelle Sprachen handelt. Aus diesem Grund ist es nur schwer möglich, dem in Gebärdensprache übersetzten Vortrag

<sup>1</sup> GESTU wird ab dem Wintersemester 2012/13 unter dem Namen „GESTU - Gehörlos erfolgreich studieren an Universitäten in Wien“ weitergeführt (siehe [Bun12]).

<sup>2</sup> In Österreich liegt die Akademikerquote bei 14,6% (siehe [Sta12]). Im Gegensatz dazu besitzen lediglich 0,1% aller gehörlosen Staatsbürger einen tertiären Bildungsabschluss (siehe [KS07, S: 397]).

zu folgen und gleichzeitig mitzuschreiben. Deshalb schreiben diese Tutoren/-innen - sehr oft Studenten/-innen, welche die Lehrveranstaltung ebenfalls absolvieren - im Kurs mit und arbeiten die Mitschrift im Anschluss an die Lehrveranstaltung gemeinsam mit den Hörbehinderten durch.

Neben diesen klassischen Unterstützungsmöglichkeiten war es aber auch Teil des Modellversuchs, die hörbehinderten Studierenden mittels technischer Hilfsmittel zu unterstützen. Hattinger beschreibt dies folgendermaßen:

„Eine offene sowie kritische Auseinandersetzung und Versuche mit verschiedensten technischen Hilfsmitteln soll dazu führen, dass neue Erkenntnisse über die verschiedenen Hilfsmittel erlangt werden und in Folge ein optimierter Einsatz und eine suggestive Verbesserung der Hilfsmittel erreicht werden kann. [...] Diese Technologien sollen bereits während aber vor allem auch nach dem GESTU Projekt hörbeeinträchtigten Studierenden zur Verfügung stehen und hörbeeinträchtigte Studenten und Studentinnen auch nach dem Modellversuch im Studienalltag unterstützen“ [Hat11, S: 5].

Als Basis dafür wurden zu Beginn des GESTU Projekts die gängigen Methoden sowie der Stand der Forschung zur Unterstützung hörbehinderter Studierender in anderen Ländern erfasst. Das wurde in der Diplomarbeit mit dem Titel „GESTU: Evaluierung von technischen Hilfsmitteln zur Förderung Studierender mit Hörbehinderung im österreichischen tertiären Bildungssektor und Einführung geeigneter Technologien an der TU Wien“ von Hattinger [Hat11] dokumentiert.

## 1.4 Motivation für diese Diplomarbeit

Die im vorigen Abschnitt beschriebene „offene sowie kritische Auseinandersetzung und Versuche mit verschiedensten technischen Hilfsmitteln“ [Hat11, S: 5] ist jedoch auch der Schwerpunkt dieser Diplomarbeit. Dafür wurden ausgehend von Hattingers Empfehlungen die vielversprechendsten technischen Maßnahmen für die Erprobung, Durchführung und Evaluierung im Zuge dieser Arbeit und während des GESTU Projekts ausgewählt. Oberstes Ziel dieser Bemühungen war die Unterstützung der hörbehinderten Studierenden und eine Erleichterung von deren Studienalltag. Dabei sollte es sich jedoch nicht, wie manchmal vermutet, um eine Bevorzugung handeln, sondern dabei „geht es um Chancengleichheit im Bildungssystem und um einen Ausgleich der Bildungsbedingungen“ [KS07, S: 369 - 370]. Beim Einsatz der technischen Maßnahmen war es wesentlich, die bestehenden klassischen Methoden zu ergänzen und entlasten.

Während der zweijährigen Laufzeit des Modellversuchs wurden die technischen Hilfsmittel gemeinsam mit den hörbehinderten Teilnehmern des Projekts angepasst und weiterentwickelt. Weiters wurde auch auf die Praxistauglichkeit, die Integration in den Regelbetrieb an den Wiener Universitäten sowie den möglichen Einsatz nach Abschluss der Diplomarbeit geachtet. Aus diesen Gründen sollten best-practice Leitfäden erstellt werden. Abschließend wurden die einzelnen eingesetzten Maßnahmen anhand der zuvor genannten Kriterien aber besonders aufgrund der Rückmeldungen der hörbehinderten Studierenden beurteilt und eine Empfehlung ausgesprochen.



**Abbildung 1.1:** Zuordnung und Abgrenzung dieser Diplomarbeit im GESTU Projekt sowie der beiden anderen im Zuge des Projekts durchgeführten Diplomarbeiten von Hattinger [Hat11] und [Hat13].

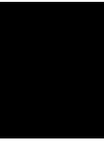
## 1.5 Abgrenzung dieser Diplomarbeit

Ziel der Diplomarbeit ist es, hörbehinderte Studierende durch technische Hilfsmittel im Studienalltag zu unterstützen. Die Unterstützung mittels klassischer Methoden erfolgte im Zuge des restlichen GESTU Projekts. Bei der Auswahl der technischen Maßnahmen wurden nur solche ausgewählt, die eine Informationsaufnahme durch den visuellen Kanal ermöglichen, unterstützen oder verbessern und somit eine Substitution des fehlenden Hörsinnes durch den Sehsinn ermöglichen oder begünstigen. Es wurden daher keine Hilfsmittel ausgewählt, die den Hörsinn verstärken - wie zum Beispiel Hörgeräte - oder einen Defekt im Hörorgan des Menschen überbrücken - wie zum Beispiel Cochlea-Implantate. Darüber hinaus wurden ebenso Maßnahmen nicht berücksichtigt, die taubblinden Personen helfen und über taktile Methoden arbeiten.

Wie bereits im vorhergehenden Abschnitt beschrieben, dient die Erfassung der gängigen Methoden zur Unterstützung hörbehinderter Studierender von Hattinger [Hat11] als Vorarbeit für diese Diplomarbeit. Daher wurde im Zuge der Arbeit keine erneute Erfassung durchgeführt, sondern ausgehend von seinen Empfehlungen die vielversprechendsten technischen Maßnahmen ausgewählt. Für die Auswahl wurde eine Sichtung der vorhandenen Forschung durchgeführt. Um die Aktualität zu gewährleisten, wurde der Suchzeitraum auf die letzten zehn Jahre beschränkt.

Der Katalog der technischen Hilfsmittel wurde in Absprache mit dem GESTU Team den am Projekt teilnehmenden Studierenden angeboten. Sie entschieden selbst, welche sie gerne probieren möchten. Da es um die Person ging, bestimmte auch sie selbst, welche Maßnahmen schlussendlich für sie geeignet waren. Infolge dessen fanden manche Hilfsmittel regen Anklang während andere kaum eingesetzt wurden. Die Versuche wurden daher hauptsächlich durch die Nachfrage und Bedürfnisse der hörbehinderten Studierenden bestimmt.

KAPITEL 2



**Untertitel**

Im vorliegenden Kapitel werden Untertitel, deren Verwendung für hörbehinderte Personen sowie die aktuell gebräuchlichen Erzeugungsmethoden behandelt. Das geschieht vor allem in Hinblick auf deren Einsatz während des GESTU Projekts und der Beschreibung im Zuge der Diplomarbeit. In weiterer Folge werden die benötigten Begriffe erläutert. Besonders die beiden nachfolgenden Kapitel über Spracherkennungssoftware<sup>1</sup> (Kapitel 3 ab Seite 17) und die Respeaking Technik (Kapitel 4 ab Seite 29) bauen darauf auf.

## 2.1 Geschichtlicher Hintergrund von Untertiteln und deren aktuelle Einsätze

Bei Untertiteln handelt es sich im Allgemeinen um schriftliche Informationen zu Bildern, Texten oder sonstigen Medien. Zum größten Teil sind sie jedoch aus dem Fernsehen und Filmen bekannt. Dort dienen sie zur schriftlichen Übersetzung des akustischen, hauptsächlich des gesprochenen Inhaltes und werden meistens am unteren Bildschirmrand eingeblendet. Sie spielen für hörbehinderte Personen eine entscheidende Rolle, um Zugang zu auditiven Informationen zu erhalten. Dabei werden sehr oft auch Zusatzinformationen wie zum Beispiel Musik, markante Geräusche oder der aktuelle Sprecher/-in schriftlich hinzugefügt. Das dient vor allem dazu um ein vollständigeres Bild des Inhaltes zu erhalten. Untertitel stellen jedoch nicht nur für schwerhörige und gehörlose Personen, sondern auch für fremdsprachige Menschen oder Sprachanfänger/-innen einen Zugang zu gesprochenen Inhalten dar.

Mit dem Aufkommen größerer TV Geräte wurden Untertitel das erste Mal in den 1970er Jahren eingesetzt, um hörbehinderten Personen die Fernsehinhalte zugänglich zu machen. Der endgültigen Durchbruch erfolgte durch die Möglichkeit der optionalen Einblendung als sogenannte „Closed Captions“. Besonders in den letzten zwei Jahrzehnten wurden die nationalen Fernsehstationen in immer mehr Ländern durch gesetzliche Regelungen dazu verpflichtet, für einen gewissen Prozentsatz ihres Programmangebotes Untertitel anzubieten. So zum Beispiel in Deutschland, den USA u.v.a. . Eine besondere Vorreiterrolle nimmt hierbei jedoch Großbritannien ein. Dort wurden bereits im Jahr 2010 alle größeren TV Sender zu einer Untertitelquote von 90% verpflichtet (siehe [Hat11, S: 47 - 50], [RF11, S: 6 - 11]). Aber auch in Österreich muss der ORF bis Ende des Jahres 2012 60% seines gesamten Angebotes mit Untertiteln anbieten (siehe [ORF12]). Infolge dieser Vorgaben sind die verschiedenen Fernsehstationen gezwungen, Untertitel in großen Mengen anzubieten beziehungsweise zu produzieren. Im TV müssen jedoch nicht nur für aufgezeichnete sondern vermehrt auch für Live-Sendungen erstellt werden. Letztere bringen einige Schwierigkeiten mit sich und stellen eine besondere Herausforderung dar. Deshalb müssen Techniken eingesetzt werden, welche die Texte mit einer möglichst kurzen Verzögerung und in ausreichender Menge und Qualität erstellen können.

Untertitel werden vermehrt auch in anderen Bereichen als dem Fernsehen oder in Filmen ein-

<sup>1</sup> In der deutschen Literatur über Spracherkennung sind fast ausschließlich die Bezeichnungen Spracherkennung, Sprachmuster, „Sprach...“ gebräuchlich. Das ist jedoch genau genommen nicht zutreffend. Es handelt sich um eine falsche Übersetzung der englischen Begriffe „speech recognition“, „speech pattern“ ... , bei denen das gesprochene Wort und nicht die Sprache im Mittelpunkt steht. Korrekt müsste es daher Sprech-Erkennung, Sprech-Muster, „Sprech...“ lauten. In der Diplomarbeit werden dennoch die Begriffe Spracherkennung, Sprachmuster,... verwendet, da sie etabliert und gebräuchlich sind (siehe [Hic13, S: 3]).

gesetzt. Zunehmend werden sie auch in Museen, Konferenzen oder an Bildungseinrichtungen verwendet. Sie dienen dort als Alternative zu (Gebärdensprach-)Dolmetscher/-innen, um hörbehinderten oder fremdsprachigen Personen gesprochene Inhalte zugänglich zu machen. Sehr oft geschieht das aus Kostengründen aber auch auf Grund mangelnder Verfügbarkeit von ausgebildeten und qualifizierten Simultanübersetzern/-innen. So können zum Beispiel für manche dieser Einsatzgebiete die Texte vorbereitet und bei Bedarf angezeigt werden. So zum Beispiel bei einem Ausstellungsstück oder Bild (siehe [RF11, S: 138 - 146]).

## 2.2 Verschiedene Arten von Untertiteln

Wie bereits zuvor angedeutet gibt es die verschiedensten Arten von Untertiteln. In diesem Abschnitt werden die wichtigsten Unterscheidungsmerkmale und Begriffsdefinitionen für die weitere Verwendung in der Diplomarbeit kurz vorgestellt.

### Untertitel für hörbehinderte Personen

Bei Untertiteln für hörbehinderte Personen wird nicht nur der gesprochene Inhalt übersetzt. Zusätzlich werden noch Informationen über den aktuellen Sprecher/-in, Musik, markante Geräusche u.a.m. gezeigt. Dadurch soll es ermöglicht werden, ein möglichst umfangreiches Bild zu erfassen (siehe [RF11, S: 17]).

### Open / Closed Caption

Bei sogenannten „Open Captions“ oder im Deutschen „offene Untertitel“ handelt es sich um dauerhaft angezeigte Untertitel. Sie können auch nicht ausgeblendet werden und wurden vor allem in deren Anfängen im Fernsehen eingesetzt. Heutzutage werden sie zum Beispiel im TV und bei fremdsprachigen Dialogen eingesetzt. Im Gegensatz dazu werden „Closed Captions“ oder „geschlossene Untertitel“ nur auf Wunsch des Betrachters/-in ausgeblendet (siehe [Hat11, S: 54], [RF11, S: 6]).

### Live, Semi-Live, Offline Untertitel

Es gibt drei verschiedenen Arten wann Untertitel erstellt werden können (siehe [Hat11, S: 50 - 51], [RF11, S: 12]):

- **Offline- oder vorbereitete Untertitel**

Bei vorbereiteten Untertiteln ist das zu untertitelnde Material bereits im Voraus bekannt, wie das bei Filmen oder bei aufgezeichneten Fernsehsendungen der Fall ist. Somit können diese vor dem eigentlichen Einsatz erstellt und anschließend mit dem Inhalt verknüpft werden. Bei dieser Methode ist es möglich, qualitativ sehr hochwertige Texte anzufertigen und zum richtigen Zeitpunkt anzuzeigen.

- **Live- oder Echtzeit-Untertitel**

Echtzeit-Untertitel werden gleichzeitig zu einem akustischen Inhalt erstellt. Dabei ist es besonders wichtig, dass sie mit einer möglichst geringen Verzögerung aber auch in ausreichender Geschwindigkeit produziert werden. Sie werden zum Beispiel im Fernsehen bei Live-Sportevents, bei Vorträgen oder in Lehrveranstaltungen eingesetzt.

- **Semi-Live Untertitel**

Hierbei handelt es sich um eine Mischform der beiden vorherigen Methoden. Vom untertitelnden Inhalt sind manche Abschnitte bereits im Voraus bekannt. Dafür werden die Untertitel vorbereitet und im richtigen Moment angezeigt. Es kommen jedoch auch im Vorhinein nicht bekannte Passagen vor, für die die Texte in Echtzeit erstellt werden müssen. Das kann bei Nachrichtensendungen mit fertigen Beiträgen und Live-Interviews der Fall sein.

### **Intralinguale oder interlinguale Untertitel**

Bei der Erstellung der Untertitel kann auch eine gleichzeitige Übersetzung in eine andere Sprache erfolgen. Ist das der Fall, spricht man von interlingualen Untertiteln. Sie werden sehr oft bei fremdsprachigen Inhalten in Filmen eingesetzt um die, oftmals dem Zuseher nicht verständlichen Dialoge zu übersetzen. Erfolgt keine simultane sprachliche Übersetzung spricht man von intralingualen Untertiteln (siehe [Hat11, S: 51], [RF11, S: 12 - 13]).

### **Wörtliche (1:1) oder gekürzte Untertitel**

Im Zuge der Erstellung der Untertitel kann der gesprochene Inhalt wörtlich (1:1) wiedergegeben werden oder aber gekürzt werden. Neben dem Zeitpunkt wann die Untertitel erstellt werden, ist das eines der wichtigsten Kriterien. Bei einer Kürzung werden inhaltlich nicht relevante Passagen zusammengefasst. Es handelt sich dabei um Lückenfüller wie ähm, ahm, uhm,... oder Hilfssätze. Das wird bei manchen manuellen Erzeugungsmethoden eingesetzt oder aber, wenn das Sprechtempo zu hoch ist.

Der aktuelle Stand der Forschung empfiehlt eine Kürzung, da das Lesetempo eines Menschen langsamer ist als gesprochene Kommunikation stattfindet. Deshalb müssen die Sätze eine bestimmte Zeit im Bild zu sehen sein. Je schneller sie wieder verschwinden, umso weniger Zeit bleibt dem/r Leser/-in zum Betrachten des visuellen Inhaltes. Im schlimmsten Fall verschwinden die Untertitel schneller, als sie gelesen werden können. Weiters belegen wissenschaftliche Forschungen, dass eine Person auch nur eine bestimmte Informationsmenge durch Lesen aufnehmen kann.

Allgemein ist das jedoch ein heikles Thema, da hierbei verschiedene Interessen aufeinander stoßen. Einige davon sind nicht auf den ersten Blick ersichtlich. Viele hörbehinderte Personen bevorzugen ungekürzte Untertitel, da sie nur so zum gesamten Inhalt Zugang haben. Eine Kürzung kommt für sie oftmals einer Zensur gleich (siehe [Hat11, S: 51 - 54], [RF11, S: 16 - 17, 112 - 116]).

## 2.3 Methoden zur Untertitelerzeugung

Zur Erzeugung von Untertiteln gibt es unterschiedliche Techniken, von denen jede einzelne bestimmte Vor- und Nachteile besitzt. Zur Unterscheidung aber auch zur Auswahl der verschiedenen Methoden gibt es vorrangig die nachfolgenden Merkmale. Manche sind ausschließlich für die Erstellung in Echtzeit maßgeblich:

- **Manuelle oder automatische Erzeugung**

Hierbei wird unterschieden, ob die Untertitel vollautomatisch von einem Computer oder von einer beziehungsweise mehreren Personen mit Hilfe von verschiedenen Eingabegeräten erstellt werden. Der Einsatz von Spracherkennungssystemen zur Erzeugung von Echtzeituntertiteln stellt diese jedoch vor große Probleme, weshalb sie in der Praxis nicht oder nur experimentell eingesetzt werden. Die prinzipielle Arbeitsweise einer Spracherkennungssoftware (ASR steht im Englischen für automatic speech recognition) aber auch die Herausforderungen beim Einsatz dieser in der Echtzeituntertitelung von gesprochenen Inhalten werden im nächsten Kapitel 3 auf Seite 17 behandelt.

- **Verzögerung**

Es handelt sich dabei um die zeitliche Verzögerung, bis die Untertitel nach dem gesprochenen Inhalt erscheinen. Das ist ausschließlich bei der Live-Untertitelerstellung wichtig.

- **Geschwindigkeit**

Bei der Geschwindigkeit handelt es sich um ein Maß dafür, wie viele Untertitel mit der jeweiligen Methode in einer gewissen Zeit erzeugt werden können. Für Echtzeituntertitel wird sie in Wörtern pro Minute (WpM) angegeben. Bei normalem Sprechtempo spricht eine Person ungefähr 140 - 160 WpM. Die Rate variiert abhängig von der Sprache selbst aber auch vom Sprechtempo. Bei schnellem Tempo kann sie auch über 220 WpM liegen (siehe [Hat13, Kapitel 2.2.2], [RF11, S: 114]).

Bei Offline-Untertiteln wird sehr oft das Verhältnis - zum Beispiel „10:1“ - angegeben, wie lange die Erstellung der Untertiteln für eine Minute des Originals in Anspruch nimmt.

- **Qualität der Untertitel**

Die Qualität von Untertiteln kann anhand verschiedener Methoden und unterschiedlicher Kriterien beurteilt werden. Sehr oft wird dabei die Wortakkuratheit angegeben. Es handelt sich dabei um die Anzahl der richtig übersetzten Wörter. Zwei Methoden der Beurteilung sind im folgenden Abschnitt 2.4 auf Seite 14 beschrieben.

- **Schwere der Fehler**

Viele der unterschiedlichen Erstellungsmethoden produzieren für sie charakteristische Fehler. Manche beeinträchtigen die Lesbarkeit der Untertitel oder verfälschen die Information stärker als andere. So verwirrt zum Beispiel ein durch eine Spracherkennungssoftware falsch erkanntes Wort oftmals den/die Leser/-in mehr als ein Tippfehler.

- **Kosten**

Es handelt sich dabei um die Kosten für den Einsatz der jeweiligen Methode. Sie hängt im Normalfall von der Ausbildungszeit der untertitelnden Person und dem eingesetzten Equipment ab.

- **Ausbildungszeit**

Die Dauer der Ausbildungszeit der Person, welche die Untertitel erstellt, hat in der Regel direkte Auswirkungen auf die Verfügbarkeit aber auch die Kosten der einzelnen Untertitelungstechniken.

Im Anschluss werden noch einige der aktuell gebräuchlichen manuellen Untertitel-Erzeugungsmethoden kurz behandelt:

### **QWERTZ Tastaturen**

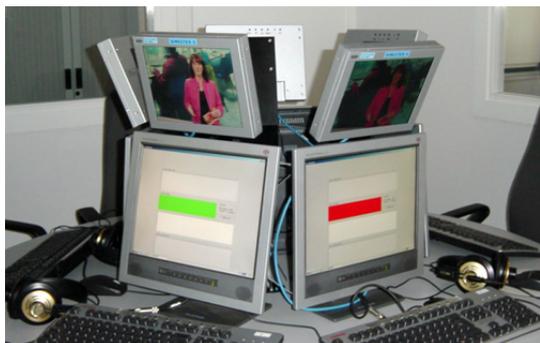
Die ersten Eingabegeräte zur Erstellung von Untertiteln waren normale Computertastaturen. Es stellte sich jedoch schnell heraus, dass sie für die Echtzeit-Untertitelung nicht ideal sind. So erreichen versierte Schnellschreiber/-innen nur ungefähr 120 WpM. Zur Verbesserung der Raten werden oftmals bis zu fünf Schnellschreiber/-innen gleichzeitig eingesetzt (siehe [Hat11, S: 77], [RF11, S: 13 - 15]). Solch ein spezieller Arbeitsplatz für mehrere Schreiber/-innen ist in Abbildung 2.1(a) zu sehen.

### **Veyboard**

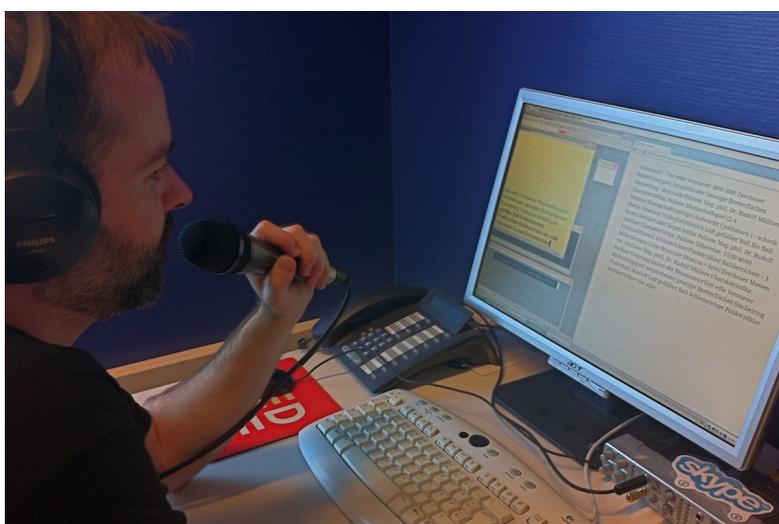
Neben dem Einsatz handelsüblicher Tastaturen kommen auch spezielle Eingabegeräte wie zum Beispiel das Veyboard zum Einsatz. Mit dessen Hilfe werden nicht nur einzelne Buchstaben sondern durch das gleichzeitige Drücken mehrerer Tasten ganze Silben und Wörter erstellt. Dabei werden durchschnittlich 120 WpM erreicht. Die Kosten für ein solches Veyboard liegen bei ungefähr 2.500 € und die Ausbildungszeit beträgt ungefähr zwölf Monate (siehe [Hat11, S: 80], [RF11, S: 13 - 15]). Ein solches Veyboard ist in Abbildung 2.1(d) zu sehen.

### **Computerstenografie**

Die klassische Stenografie selbst blickt auf eine sehr lange Geschichte zurück und kommt heute in Österreich hauptsächlich bei der Mitschrift in Parlamentssitzungen zum Einsatz. Es handelt sich dabei jedoch um keine Methode zur Untertitelerzeugung. Erst unter Zuhilfenahme spezieller Eingabegeräte in Verbindung mit einem Computer kann sie auch zur Untertitelung eingesetzt werden. Wie auch beim Veyboard werden mehrere Tasten gleichzeitig gedrückt, um so ganze Silben und Wörter zu erstellen. Dadurch werden sehr hohe Geschwindigkeiten von 220 bis zu maximal 300 WpM erreicht. Die Kosten für ein Computerstenografiegerät betragen ungefähr 1.800 - 4.000 € und die Ausbildungszeit ist mit drei Jahren sehr lange. Die Methode wird bevorzugt in den USA eingesetzt, findet in Österreich jedoch kaum Anwendung (siehe [Hat11, S: 78 - 79], [RF11, S: 13 - 15]). Ein Computerstenograf ist in Abbildung 2.1(b) zu sehen.



(a) Ein System mit dem mehrere Schnellschreiber zusammenarbeiten um mithilfe von QWERTZ Tastaturen Untertitel zu erstellen (b) Eine Stenografin bei der Arbeit mit einem Computerstenografiesystem



(c) Ein Respeaker der Firma Titelbild



(d) Ein Veyboard

**Abbildung 2.1:** Verschiedene Methoden zur manuellen Untertitel-Erzeugung [Hat 11, Abbildungen 3.7 und 3.9].

## Respeaking

Eine der jüngsten, aber in den letzten Jahren stark zunehmende Methode ist das Respeaking. Dabei spricht eine Person, ein/-e sogenannte/-r Respeaker/-in den gesprochenen Inhalt nochmals für eine Spacherkennungssoftware nach. Sie wird dabei, wie zuvor die verschiedenen Tastaturen, als Eingabemethode verwendet. Die Respeaking Technik wird im Kapitel 4 auf Seite 29) genauer beschrieben. Die Vorteile gegenüber den zuvor genannten Methoden sind hohe Geschwindigkeiten von 160 - 190 WpM bei vergleichsweise geringen Ausbildungszeiten von wenigen Monaten. In Abbildung 2.1(c) ist ein Respeaker der Firma Titelbild zu sehen.

Abschließend eine Gegenüberstellung der zuvor beschriebenen Techniken anhand der am Beginn dieses Abschnitts beschriebenen Kriterien:

|                            | <b>Verzögerung</b> | <b>Geschwindigkeit (WpM)</b>    | <b>Wortakkuratheit</b> | <b>Ausbildungszeit</b> | <b>Kosten</b> | <b>Schwere der Fehler</b> |
|----------------------------|--------------------|---------------------------------|------------------------|------------------------|---------------|---------------------------|
| <b>Keyboard</b>            | mittel             | mittel<br>(90 - 120)            | 95%                    | 12<br>Monate           | mittel        | hoch                      |
| <b>Dual Tastatur</b>       | mittel             | mittel bis<br>hoch (140 - 150)  | 95 - 98%               | 6<br>Monate            | mittel        | niedrig                   |
| <b>Computerstenografie</b> | niedrig            | sehr hoch<br>(220 - bis zu 300) | 97 - 98%               | 3<br>Jahre             | hoch          | mittel                    |
| <b>Respeaking</b>          | niedrig            | hoch<br>(160 - 190)             | 97 - 98%               | 2 - 3<br>Monate        | niedrig       | mittel -<br>hoch          |

**Tabelle 2.1:** Vergleich manueller Live-Untertitelungsmethoden [RF11, Table 2.1].

## 2.4 Qualität von Untertiteln bestimmen

Bei der Erzeugung von Untertiteln spielt nicht nur die Geschwindigkeit eine große Rolle, sondern auch deren Qualität. Zur Bestimmung gibt es verschiedene Ansätze und Kriterien. Eine der ältesten, bewährtesten und auch einfach anzuwendenden Methoden ist die Bestimmung der Wortfehlerrate (WER steht im Englischen für word error rate) beziehungsweise der Wortakkuratheit (WRR steht im Englischen für word recognition rate). Die beiden Parameter stammen aus dem Bereich der Spacherkennungssoftware. Dort sind sie Standard zur Angabe der Qualität solcher Systeme beziehungsweise der Transkripte (siehe [Hat11, S: 63 - 64], [RF11, S: 150 - 152]).

Bei der WER wird die Summe der falsch erkannten Wörter S (steht im Englischen für Substitutions), die Anzahl der falsch eingefügten Wörter I (steht im Englischen für Insertions) und die Menge der nicht erfassten Wörter D (steht im Englischen für Deletions) der Gesamtanzahl aller Wörter N (steht im Englischen für Number of words) gegenübergestellt:

$$WER = \frac{D + S + I}{N} \times 100 = [\%] \quad (2.1)$$

Sehr oft wird auch die Wortakkuratheit verwendet, für die Werte von über 97% sehr gut sind. Sie wird folgendermaßen berechnet:

$$WRR = 1 - WER = \frac{N - D - S - I}{N} \times 100 = [\%] \quad (2.2)$$

Mit der WER kann die Qualität von ungekürzten Untertiteln relativ schnell und einfach bestimmt werden. Sie hat jedoch auch Nachteile. Es wird die zuvor beschriebene Schwere der Fehler, also deren Auswirkung auf die Lesbarkeit beziehungsweise die Information in der Berechnung nicht berücksichtigt. Sie spielt jedoch mitunter eine große Rolle. Manche Fehler können vom/von der Leser/-in selbstständig mithilfe des Kontexts korrigiert werden, wohingegen andere die Bedeutung des Inhaltes komplett verändern (siehe [RFM13], [Wal06]).

Des Weiteren eignet sich die Wortfehlerrate nicht zur Bewertung von gekürzten Texten. Sie wertet korrekte Zusammenfassungen und Umformulierungen als Fehler und verfälscht somit das Ergebnis. Aufgrund dessen, aber unter Verwendung des Prinzips der WER, entwickelte Romero-Fresco und Juan Martínez das sogenannte NER Modell (siehe [RFM13]). Es wurde besonders für die Qualitätsbestimmung von mithilfe der Respeaking Technik erstellten Untertitel entwickelt. Diese Methode berücksichtigt noch verschiedene andere Faktoren. Dennoch lässt sich das Modell für sämtliche Transkripte, unabhängig von deren Erstellungsmethode, anwenden. Die Berechnung funktioniert dabei folgendermaßen (Ergebnisse größer 98% sind als gut zu bewerten):

$$\text{Genauigkeit} = NER = \frac{N - E - R}{N} \times 100 = [\%] \quad (2.3)$$

- **N: Anzahl der Wörter im transkribierten Text** (im Englischen **N**umber of words in the respoken text):  
Das ist die Gesamtanzahl der Wörter im transkribierten Text. Dazu zählen auch Interpunktionszeichen, Sprecherinformationen, ... .
- **E: Editierfehler** (im Englischen **E**dition Errors):  
Bei Editierfehlern handelt es sich um Fehler durch eine falsche Bearbeitung. Zur Feststellung und Bewertung wird der transkribierte mit dem originalen Text verglichen. Bei gekürzten Untertiteln handelt es sich um nicht korrekte Kürzungen oder Zusammenfassungen. Bei ungekürzten Transkripten sind das zum Beispiel eine falsche Groß- oder Kleinschreibung oder falsche Satzzeichen.
- **R: Erkennungsfehler** (im Englischen **R**ecognition Errors):  
Es handelt sich dabei um falsch erkannte Wörter, verursacht durch die eingesetzte Eingabemethode. Zum Beispiel können das durch die Spracherkennungssoftware hinzugefügte, ausgelassene oder nicht richtig erkannte Wörter sein. Bei den Erkennungsfehlern werden alle Fehlerarten des WER Modells zusammengefasst. Sie werden ebenfalls durch einen Vergleich des Originals mit dem Transkript bestimmt und bewertet.

Die Editier- und die Erkennungsfehler werden anhand der folgenden Kriterien beurteilt und gewichtet:

- **Gravierende Fehler**

Sie führen zur einer Bedeutungsänderung des Inhaltes, der aber im Kontext des Originaltextes Sinn ergeben kann. Gravierende Fehler werden mit einer Gewichtung von 1 miteinberechnet. Als Beispiel nennt Romero-Fresco [RFM13] die Transkription von „in Island ist der Zinssatz von 3,5 auf 2% gesunken“ anstatt „Irland ist der Zinssatz von 3,5 auf 2% gesunken“.

- **Normale Fehler**

Normale Fehler führen im Gegensatz zu gravierenden Fehlern nicht zu einer neuen Bedeutung des Inhaltes. Dennoch kommt es zu einem Informationsverlust und sie stören den Lesefluss. Sie werden mit 0,5 in der Berechnung berücksichtigt. Als Beispiel nennt Romero-Fresco [RFM13] die Transkription von „die Chancen werden immer kleiner, die Grünen haben ihren Ausdruck beschlossen“ anstatt „die Grünen haben ihren Austritt beschlossen“.

- **Geringfügige Fehler**

Diese Fehler verfälschen weder die Information des Originaltextes noch unterbrechen sie den Lesefluss. In manchen Fällen kann es sogar möglich sein, dass der/die Leser/-in sie mithilfe des Kontexts selbstständig korrigieren kann. Geringfügige Fehler werden mit einem Wert von 0,25 in der Berechnung gewichtet. Als Beispiel nennt Romero-Fresco [RFM13] die Transkription von „dass die irische Regierung zwei weitere konkursreifer Banken Verstaatlichung muss“ anstatt „dass die irische Regierung zwei weitere konkursreife Banken verstaatlichen muss.“.

Zur Komplettierung des NER Modells werden noch zwei weitere Parameter hinzugefügt. Sie fließen nicht in die zuvor beschriebene Berechnung mit ein, helfen jedoch bei der Beurteilung der Untertitel:

- **Korrekte Änderungen** (im Englischen **Correct Editions**)

Es handelt sich dabei um korrekte Kürzungen und Zusammenfassungen. Sie verändern den Inhalt des Textes nicht (siehe [RFM13]).

- **Beurteilung** (im Englischen **Assessment**)

In der Beurteilung werden abschließend zuvor nicht erfasste Kriterien, wie zum Beispiel die Verzögerung der Untertitel, deren Anzeige und Zusammenhang mit den Bildern beurteilt [RFM13]).

Durch die Klassifizierung und Gewichtung der Fehler ist eine Beurteilung der Qualität von Untertiteln anhand des NER Modells aufwändiger durchzuführen. Das Ergebnis ist von der subjektiven Einschätzung der beurteilenden Person abhängig. Dadurch erfüllt diese Methode auch die auf Seite 15 geforderte Berücksichtigung der Auswirkung von einzelnen Fehlern. Somit ist eine aussagekräftigere Bewertung von Transkripten möglich.

KAPITEL 3

**Spracherkennungssoftware**

Wie bereits zuvor erwähnt gibt es verschiedene Methoden, um Untertitel unter Zuhilfenahme von Spracherkennungssoftware<sup>1</sup> (ASR steht im Englischen für automatic speech recognition) zu erzeugen. In diesem Kapitel wird die prinzipielle Funktionsweise von solchen Systemen sowie deren Einsatz zur automatischen Erstellung von Transkripten erläutert. Der zweite Teil des Abschnitts beschreibt das Training für den österreichischen Dialekt eines sprecherunabhängigen Spracherkennungssystems für freie Sprache sowie den Einsatz im Zuge des GESTU Projekts.

### 3.1 Stand der Technik und Forschung

Seit dem Aufkommen der modernen Computertechnik besteht das Bestreben, gesprochene Sprache automatisch und in Echtzeit durch Spracherkennungssysteme zu transkribieren. Das erwies sich nach der anfänglichen Euphorie jedoch schnell als nicht so einfach lösbar. Heute wird die Lage in Fachkreisen wesentlich realistischer und nüchterner beurteilt. Trotzdem existiert in der Bevölkerung zum Teil immer noch kein klares Bild über die Leistungsfähigkeit aktueller Systeme beziehungsweise über deren Grenzen. Es existiert die Meinung, dass die Untertitel im Fernsehen durch Spracherkennungssoftware ohne menschliches Zutun erstellt werden. In der über 60 jährigen Geschichte der Forschung und Entwicklung auf dem Gebiet der Spracherkennungssysteme und der Fortschritte in den letzten Jahrzehnten gibt es immer noch keine allgemeine Lösung für des Problem (siehe [PK08, S: 21, 283 - 300], [RF11, S: 56 - 71, 173]). Kaufmann und Pfister kommen weiters zu der Erkenntnis, „dass trotz grosser Forschungsanstrengungen die maschinelle Spracherkennung auch in 20 Jahren noch nicht den Stand der menschlichen Sprachwahrnehmungsfähigkeit erreicht haben wird“ [PK08, S: 21].

Die Leichtigkeit, mit der wir Menschen Sprache nicht nur verstehen sondern auch die Bedeutung des Gesagten erfassen, täuscht über die eigentliche Komplexität der Aufgabenstellung hinweg. Für ein Spracherkennungssystem, dessen Ziel es ist, nur die gesprochenen Wörter zu erkennen, ohne den Inhalt zu verstehen, stellt alleine das schon ein sehr großes Problem dar.

Die Schwierigkeit beim Erkennen einzelner Laute und in weiterer Folge von Wörtern und Sätzen entspringt aus der Natur der gesprochenen Sprache und deren Erzeugung im menschlichen Stimmapparat, sowie der sich daraus ergebenden großen Variabilität eines Sprachsignals. Es ist ein und der selben Person praktisch unmöglich, etwas zweimal so zu sagen, dass die Sprachsignale identisch sind. Die Unterschiede sind für verschiedene Sprecher/-innen noch bei weitem größer. Das ergibt sich aus den körperlichen Voraussetzungen des/der Sprechers/-in. Es handelt sich dabei um die Beschaffenheit des Vokaltraktes, das Geschlecht, das Alter und den Gesundheits- sowie den Emotionszustand. Weiters spielt noch die Herkunft und Sprechweise eine Rolle. Von diesen sprecherabhängigen Einflussfaktoren abgesehen ist die Aussprache eines Lautes von den Nachbarlauten abhängig. Das wird die Koartikulation genannt. Wird zum Beispiel der Buchstabe „k“ gemeinsam mit einem „i“ ausgesprochen, klingt er anders als in Kombination mit einem „u“. Zuletzt hat auch der Übertragungsweg, aber auch die Geräuschkulisse, noch einen wesentlichen Einfluss auf das Sprachsignal (siehe [PK08, S: 21, 283 - 300]). Die verschiedenen Einflussfaktoren auf ein Sprachsignal sind in der nachfolgenden Abbildung 3.1 dargestellt.

---

<sup>1</sup> An der Stelle sei nochmals auf die Erklärung bezüglich der, im Deutschen üblichen Begriffe Spracherkennung, „Sprach...“ versus Sprecherkennung, „Sprech...“ in der Fußnote auf Seite 8 hingewiesen.

Dunkelgrau sind die sprecherabhängigen Faktoren und hellgrau die Beeinflussung durch die Übertragung des Signals selbst gekennzeichnet.

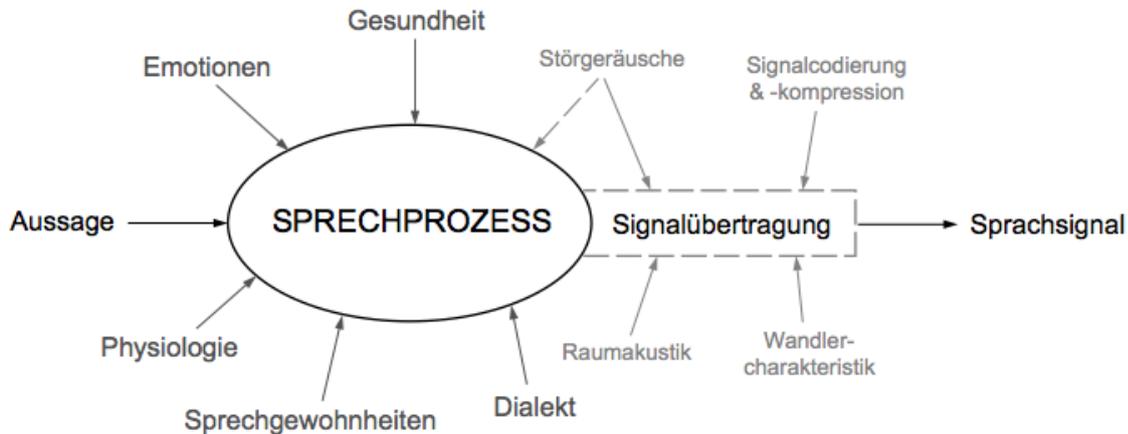


Abbildung 3.1: Einflussfaktoren auf ein Sprachsignal [PK08, Abbildung 2.1].

### 3.1.1 Unterteilung von Spracherkennungssystemen

Um die Komplexität des Problems zu vereinfachen und somit die Leistung einer Spracherkennungssoftware und damit in weiterer Folge auch die Genauigkeit zu verbessern, wird diese auf bestimmte Anwendungsfälle spezialisiert. Es kann sich dabei um die Erkennung einzelner gesprochenen Wörter, ein vorhergehendes Training durch eine Person, Erkennungssysteme für Telefonsignale oder Diktiersysteme handeln.

„Mit dieser Spezialisierung erreicht man zweierlei: Erstens zeigt sich, dass bessere Resultate erreicht werden können (weniger Erkennungsfehler) und zweitens sind diese Speziallösungen kompakter (weniger Rechenleistung und geringerer Speicheraufwand erforderlich) als eine allgemeinere Lösung und damit wirtschaftlicher“ [PK08, S: 28]

Es gibt daher für aktuelle Spracherkennungssysteme verschiedene Einsatzbereiche und Unterscheidungskriterien. Sie können zum Beispiel nach dem Aufgabengebiet unterteilt werden. Laut Pfister und Kaufman (siehe [PK08, S: 291]) gibt es folgende Einsatzgebiete:

- **Einzelworterkenner** erkennen, wie der Name bereits vermuten lässt, einzelne durch Pausen getrennte Wörter. Solche Systeme werden in der Regel auf ein beschränktes Vokabular trainiert. Ein typisches Beispiel sind Sprachsteuerungen.
- **Keyword-Spotter** bestimmen aus fließend gesprochenen Sätzen einzelne Wörter oder Kommandos. Sie können ebenfalls für Sprachsteuerungen verwendet werden.
- **Verbundworterkenner** erkennen wenige gesprochene Wörter aus einem begrenztem Vokabular. Sie werden gemeinsam ausgesprochen, haben jedoch meist keinen Zusammen-

hang. Das wird bei der Erkennung von Ziffernfolgen - zum Beispiel Telefonvorwahlen - eingesetzt.

- **Kontinuierliche Spracherkennung** wandeln gesprochene Sprache in einen geschriebenen Text um. Die einzelnen Wörter werden dabei ohne extra Pausen ausgesprochen. Bei diesen Systemen kommen in der Regel sehr große Vokabulare zum Einsatz, die sehr leicht erweiterbar sind. Um hierbei die Erkennungsraten zu verbessern, werden sehr oft auch die Gesetzmäßigkeiten der entsprechenden Sprache berücksichtigt. Bei heute üblichen Diktiersystemen handelt es sich um solche Spracherkennungssoftware.

Ein weiteres Unterscheidungskriterium für Spracherkennungssysteme ist, ob sie vor oder während des Gebrauchs vom/von der Sprecher/-in trainiert werden müssen, um somit bessere Erkennungsraten zu erzielen (siehe [Hat11, S: 61 - 62], [PK08, S: 292]):

- **Sprecherabhängige Systeme** werden auf eine/-n Sprecher/-in trainiert. Das erfolgt meist vor dem ersten Einsatz, indem ein kurzer Text vorgelesen werden muss. Dadurch wird die Spracherkennungssoftware auf den/die Sprecher/-in abgestimmt und erzielt somit bessere Ergebnisse.
- **Sprecherunabhängige Systeme** können für beliebige Sprecher/-innen ohne vorhergehendes Training eingesetzt werden. Aufgrund dieser Tatsache müssen sie meist unter schwierigeren Rahmenbedingungen operieren. Sie erzielen in der Regel schlechtere Ergebnisse als sprecherabhängige Spracherkennungssoftware und benötigen bei der Entwicklung eine ungleich größere Menge an Sprachproben.
- **Sprecheradaptive Systeme** sind prinzipiell sprecherunabhängige Systeme, können jedoch zusätzlich auf bestimmte Sprecher/-innen trainiert werden. Das kann entweder im Zuge des Vorlesens eines kurzen Trainingstextes aber auch während des Einsatzes selbst erfolgen.

Neben den oben angeführten Unterscheidungskriterien gibt es noch einige weitere - wie zum Beispiel die Vokabelgröße (siehe [PK08, S: 290 - 292]). Abschließend werden noch zwei weitere behandelt, die besonders im Verlauf dieser Diplomarbeit wichtig sind. Es handelt sich um die Verzögerung, mit der der erkannte Text zur Verfügung gestellt wird (siehe [Hat11, S: 63]):

- **Live oder Echtzeit Transkription**  
Bei solchen Systemen wird der gerade gesprochene Inhalt sofort verarbeitet. Somit stellen sie die erkannten Wörter innerhalb weniger Sekunden bereit. Hierbei ist vor allem die Verzögerung bis zur Anzeige eine wichtige Kenngröße.
- **Offline Transkription**  
Offline ASRs stellen das Ergebnis erst nach einer mitunter längeren Wartezeit bereit. Sie benötigen oftmals länger als eine halbe Minute, bis zu Stunden, und eignen sich daher nicht mehr zur Echtzeittranskription.

Weiters gibt es speziell bei kontinuierlichen Spracherkennungssystemen noch die Unterscheidung, ob sie zum Diktieren oder für freie, spontane Sprache ausgelegt sind (siehe [Hat11, S: 62]):

- **Spracherkennungssoftware zum Diktieren** wird speziell als alternative Eingabemethode für Computer entwickelt. Aus dem Grund werden sie darauf spezialisiert, vollständige und grammatikalisch korrekte Sätze zu verarbeiten. Dabei ist auch die explizite Ansage von Satzzeichen - zum Beispiel „Beistrich“ oder „Punkt“ - erforderlich. Solche Systeme müssen fast immer auf eine/-n Sprecher/-in trainiert werden. Erfolgt das alles nicht, erzielen diese ASRs schlechte und oftmals kaum oder nur schwer lesbare Ergebnisse. Darüber hinaus sind sie für die Verwendung in Büroumgebungen entwickelt und reagieren empfindlich auf Umgebungsgeräusche.
- **Spracherkennungssysteme für freie, spontane Sprache** werden zur Transkription von Sprache im Alltagsgebrauch eingesetzt. Dabei handelt es sich jedoch um ein wesentlich komplexeres und schwierigeres Anwendungsgebiet. Die Systeme müssen unter erheblich schwierigeren Rahmenbedingungen arbeiten, die sich oftmals nicht so leicht kontrollieren lassen. Bei freier Kommunikation kommt es zu Auslassungen, Unterbrechungen, Pausen und grammatikalisch inkorrekten Sätzen. Außerdem werden im normalen Sprachgebrauch keine Satzzeichen ausgesprochen. Daher sind die einzelnen Satzteile beziehungsweise Sätze wesentlich schwieriger voneinander zu trennen.

### 3.1.2 Lösungsansätze für Spracherkennungssysteme

Nachdem nun die Schwierigkeiten, die bei der Entwicklung von Spracherkennungssoftware zu lösen sind, sowie die daraus resultierende Spezialisierung erörtert wurde, folgt nun ein kurzer Überblick über die aktuell verwendeten Lösungsansätze. Das Hauptaugenmerk liegt bei kontinuierlichen Erkennungssystemen mit großen Wörterbüchern, da sie beim Respeaking zum Einsatz kommen. Außerdem handelt es sich bei der im Zuge der Diplomarbeit trainierten ASR um solch ein System. Die Durchführung und die Ergebnisse werden im zweiten Abschnitt des Kapitels ab Seite 24 beschrieben.

Abhängig vom Anwendungsgebiet werden laut Pfister und Kaufmann (siehe [PK08, S: 304 - 366]) heutzutage zwei grundlegende Lösungsansätze für Spracherkennungssysteme verwendet:

#### **Spracherkennung mittels Mustervergleich**

Hierbei wird für jedes zu erkennende Wort vorab ein Sprachmuster gespeichert. Die vorhandenen Muster werden mit dem der zu erkennenden Sprachprobe verglichen. Das Ergebnis wird anhand der höchsten Übereinstimmung ermittelt. Dieser Ansatz wird praktisch nur zur Erkennung einzelner Wörter und für ein begrenztes Vokabular verwendet (siehe [PK08, S: 304 - 323]).

#### **Statistische Spracherkennung**

Der statistische Ansatz zur Spracherkennung ermittelt durch Wahrscheinlichkeiten das am besten zu einer zu erkennenden Sprachprobe passende Ergebnis. Das geschieht anhand von ausgewählten Merkmalen des Sprachsignals. Hierbei wird nun eine Wortfolge so gewählt, dass die Wahrscheinlichkeit für eine Fehlentscheidung möglichst gering ist. Um die Entscheidung zu treffen, werden zwei Wahrscheinlichkeiten berechnet, die in zwei Modellen abgebildet werden (siehe [PK08, S: 324 - 366]):

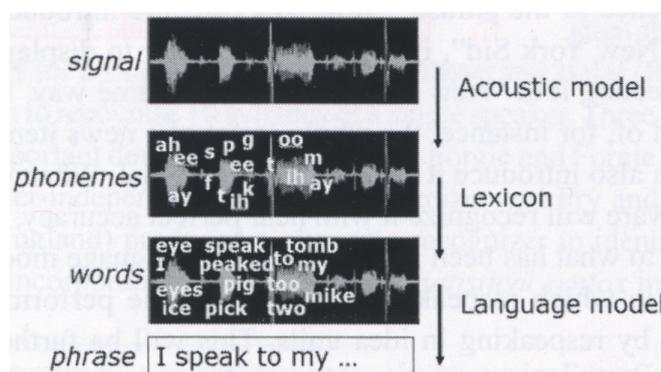
- **Akustische Modell**

Im akustischen Modell wird die Wahrscheinlichkeit abgebildet, mit der Merkmale eines Sprachsignals bei einer gewissen Lautfolge zu beobachten sind. Dadurch ist es möglich, eine Lautfolge zu wählen, die am wahrscheinlichsten zu der zu erkennenden gesprochenen Äußerung passt. Das akustische Modell wird anhand vieler verschiedener Sprachproben trainiert. Mithilfe dieser werden die Wahrscheinlichkeiten für das Auftreten von Merkmalen bei bestimmten Lautfolgen errechnet und gespeichert. Sie sollten unter den selben Rahmenbedingungen aufgezeichnet werden, wie sie später beim Einsatz des Spracherkennungssystems vorherrschen. Im Allgemeinen gilt, dass mehr Trainingsdaten später zu besseren Ergebnissen führen. Daher sollten jedoch mindestens zehn Beobachtungen pro zu erkennendem Parameter vorliegen. Für eine sprecherunabhängige ASR sind dabei mehr Proben notwendig (siehe [PK08, S: 324 - 366]).

- **Sprachmodell**

Das Sprachmodell speichert und liefert die Wahrscheinlichkeiten, mit der bestimmte Buchstaben- und Wortfolgen auftreten. Es umfasst ein Lexikon, in dem die verschiedenen Vokabel gespeichert sind. In komplexeren Systemen fließt darüber hinaus auch noch linguistisches Wissen mit ein. Es handelt sich zum Beispiel um die Wahrscheinlichkeit für das Auftreten bestimmter Wortfolgen. Das ist unabhängig von den Merkmalen des gerade zu erkennenden Sprachsignals und bereits im Voraus bekannt. Es ist von verschiedenen Faktoren abhängig, wie unter anderem der Kommunikationssituation und der daraus resultierenden Wortwahl aber auch vom Wissen über die Sprache allgemein (siehe [PK08, S: 324 - 366]).

Zum Abschluss ist noch der prinzipielle Ablauf eines Spracherkennungsvorganges in der nachfolgenden Abbildung 3.2 zu sehen. Anhand verschiedener Merkmale des Sprachsignals werden unter Zuhilfenahme des akustischen Modells die einzelnen Laute bestimmt. Sie werden im Sprachmodell zu Worten und Wortfolgen umgewandelt und abschließend wird das Ergebnis erstellt.



**Abbildung 3.2:** Schematischer Ablauf einer Spracherkennung [RF11, Figure 5.1].

### 3.1.3 Überblick über aktuelle Spracherkennungssoftware

Abschließend folgt ein kurzer Überblick über einige der aktuell in den verschiedenen Bereichen eingesetzten Spracherkennungssysteme:

#### Dragon NaturallySpeaking

Dragon NaturallySpeaking (DNS) des Herstellers Nuance<sup>2</sup> wurde 2012 in der Version 12 veröffentlicht. Es handelt sich um ein sprecherabhängiges Diktiersystem für kontinuierliche Sprache und ist unter anderem für Deutsch, Englisch, Französisch und Spanisch verfügbar. Laut Angaben des Herstellers lässt sich eine Genauigkeit von 99% erzielen. DNS analysiert außerdem je nach verwendeter Version sowie verwendetem Sprachmodell bis zu vier vorangegangene und nachfolgende Wörter. Dadurch erscheinen die erkannten Wörter in Blöcken. Das ist besonders beim Respeaking zu berücksichtigen, um die Verzögerung möglichst kurz zu halten (siehe [Hat11, S: 66 - 69], [RF11, S: 66 - 68]).

#### ViaVoice

ViaVoice ist ein von IBM<sup>3</sup> entwickeltes sprecherabhängiges Diktiersystem für kontinuierliche Sprache. IBM hat die Vertriebsrechte an Nuance verkauft und daraufhin wurde die Entwicklung eingestellt. Die letzte Version wurde 2003 veröffentlicht und ist unter anderem für Deutsch, Englisch, Französisch und Spanisch verfügbar. Die Software ist dennoch in manchen Ländern noch käuflich erhältlich und steht außerdem den Mitgliedern des Liberated Learning Consortiums (LLC)<sup>4</sup> zur Verfügung (siehe [Net07]).

Daher wird ViaVoice noch von manchen Firmen zur Erzeugung von Untertiteln mithilfe der Respeaking Technik eingesetzt. Das hat vor allem den Grund, dass die Software sehr schnell ist und die erkannten Wörter sehr schnell anzeigt. Bei entsprechendem Training erreicht das System normal 95 - 99% Genauigkeit (siehe [Hat11, S: 66 - 69], [RF11, S: 65 - 66]).

#### Spracherkennung des Liberated Learning Consortiums

Beim Liberated Learning Consortium (LLC) handelt es sich um einen Zusammenschluss von 21 Universitäten. Deren Ziel ist es, barrierefreies Lernen mithilfe von automatischen Spracherkennungssystemen zu ermöglichen. Durch den Zusammenschluss wurde unter anderem auch ein sprecherunabhängiges System für freie Sprache entwickelt. Mithilfe des „IBM Hosted Transcription Service Engine Controllers“ ist es den Mitgliedern des LLC möglich, offline Untertitel für die englische Sprache zu erstellen. Es werden Erkennungsraten von ~78,6% erzielt (siehe [Hat11, S: 66 - 69], [RF11, S: 63 - 70]).

<sup>2</sup> <http://www.nuance.de> - letzter Zugriff am 06.03.2013.

<sup>3</sup> <http://www.ibm.com> - letzter Zugriff am 06.03.2013.

<sup>4</sup> <http://liberatedlearning.com> - letzter Zugriff am 06.03.2013.

## Google

Seit 2010 bietet Google ebenfalls ein sprecherunabhängiges Spracherkennungssystem an. Es erlaubt Benutzern/-innen der Videoplattform YouTube<sup>5</sup> für ihre hochgeladenen Videos automatisch Untertitel zu erstellen. Diese Funktion war lange Zeit nur für die englische Sprache verfügbar und erreicht eine Genauigkeit von ~55,6%. Seit Ende November 2012 und nach dem Abschluss des praktischen Teiles der Diplomarbeit, ist die ASR auch unter anderem für die deutsche Sprache verfügbar (siehe [Hat11, S: 66 - 69], [RF11, S: 69 - 70], [You12]).

## Spracherkennung des European Media Labors

Zu Beginn des Net4Voice Projekts (2007) wurde erhoben, dass damals keine sprecherunabhängige automatische Spracherkennung für freie deutsche Sprache und deren Einsatz zur Transkription von Vorlesungen existierte. Während des Projekts wurde daher von der beteiligten Universität Ulm eine Kooperation mit dem European Media Labor (EML) zur Entwicklung eines solchen Spracherkennungssystems gestartet. Bereits beim Abschluss des Net4Voice Projekts (2010) stand eine funktionsfähige ASR zur Verfügung. Sie erzielte für die hochdeutsche Sprache während einer Demonstration bei dem im nächsten Abschnitt beschriebenen Treffen am 27.09.2010, bereits gute Ergebnisse<sup>6</sup> (siehe [TGN<sup>+</sup>10], [Hat11, S: 66 - 69, 91]).

## 3.2 Training der EML ASR im Zuge des GESTU Projekts

Hattinger knüpfte während seines Besuches der Net4Voice Abschlusskonferenz im Zuge der Recherche für die Diplomarbeit [Hat11] Kontakte zu den Entwicklern der zuvor beschriebenen Spracherkennungssoftware des European Media Labors (siehe [Hat13, Kapitel 1.5.2]). Ausgehend davon kam es am 27. September 2010 in Heidelberg zu einem Treffen mit den Verantwortlichen des EML, gemeinsam mit den Kooperationspartnern der Universität Ulm sowie Christian Hattinger, Prof. Dr. Wolfgang Zagler und dem Autor der vorliegenden Arbeit. Das Ziel war, mögliche Einsatzgebiete dieser Spracherkennung für die Unterstützung hörbehinderter Studierender im österreichischen tertiären Bildungssektor im Zuge des GESTU Projekts zu erörtern. Die frühe Zusammenarbeit soll auch eine längerfristige, kostengünstige Nutzung der Spracherkennung für GESTU eröffnen.

Mit dem Spracherkennungssystem des EML, nachfolgend kurz als EML ASR bezeichnet, ist es möglich, Audiodateien automatisiert zu transkribieren. Aktuell ist es nicht möglich mit dem System Untertitel in Echtzeit zu erstellen. Die erstellten Transkripte werden mit Zeitcodes versehen bereitgestellt. Dadurch ist die Nutzung im Anschluss an einen Vorlesungstermin in Kombination mit der in Abschnitt 5.2 auf Seite 57 beschriebenen Plattform Synote möglich. Im Erfolgsfall sollte es somit möglich sein, für Lehrveranstaltungen Untertitel automatisch zu erstellen und durch Synote gemeinsam mit der Videoaufzeichnung und der Folienpräsentation den am GESTU Projekt teilnehmenden hörbehinderten Studierenden zur Verfügung zu stellen.

<sup>5</sup> <http://www.youtube.com> - letzter Zugriff am 06.03.2013.

<sup>6</sup> Dabei handelte es sich um einen subjektiven Eindruck seitens des Autors. Dieser entstand während des am 27.10.2010 stattgefundenen Treffens mit den Entwicklern der EML ASR. Im Zuge dessen wurden jedoch keine genauen Zahlen über die Erkennungsdaten präsentiert.

Mit den beschriebenen Zielen im Hintergrund wurden während des Treffens alle notwendigen Schritte zur Anpassung und dem Einsatz der EML ASR im Zuge des GESTU Projekts besprochen. Bis zum damaligen Zeitpunkt wurde das System nur unter Zuhilfenahme von österreichischen Nachrichtensendungen trainiert und verwendet. Im österreichischen Fernsehen wird jedoch hauptsächlich hochdeutsch gesprochen. Dies unterscheidet sich zum Teil maßgeblich vom im Alltag gesprochenen, österreichischen Dialekt sowohl im Satzbau, dem verwendeten Vokabular und der Aussprache. Ein weiteres Problem stellen die je nach wissenschaftlicher Fachrichtung unterschiedlichen Fachausdrücke dar. Resultierend daraus ist ein Training des Systems mit einer großen Menge verschiedener Audioaufzeichnungen von Vorlesungen notwendig. Dadurch sollen mithilfe der ASR möglichst gute Erkennungsraten zur Transkription von Lehrveranstaltungen erzielt werden.

### 3.2.1 Durchführung

Ausgehend von den, im vorhergehenden Abschnitt beschriebenen Zielen, wurde mit der Sammlung von Trainingsdaten begonnen. Aufgrund des geplanten Einsatzgebietes des Spracherkennungssystems zur Transkription von Vorlesungen handelte es sich dabei um die Videoaufzeichnung von Lehrveranstaltungen. Um ein möglichst breites Spektrum an Trainingsdaten zu erhalten, wurde versucht, möglichst viele Kurse, die von am GESTU Projekt teilnehmenden hörbehinderten Studierenden besucht wurden, aufzuzeichnen. Die Durchführung der Vorlesungsaufzeichnungen ist in Abschnitt 5.3 auf Seite 59 detailliert beschrieben. Hierbei wurde auf die gute Verständlichkeit der Übersetzung der Gebärdensprachdolmetscher/-innen geachtet. Der gesprochene Vortrag wurde in möglichst guter Tonqualität aufgezeichnet und zum Training des Spracherkennungssystems verwendet. Die Tonspur wurde anschließend aus dem Video extrahiert. Da die Aufnahmen im laufenden Lehrveranstaltungsbetrieb stattfanden, ist deren Qualität für die Verarbeitung durch eine Spracherkennungssoftware nicht optimal. Das ist durch wechselnde Sprechlautstärke im Vortrag, unterschiedliche Aussprachen und Nebengeräusche bedingt und beeinträchtigt die Qualität des Ergebnisses.

Bevor die Tonspur der Vorlesungsaufzeichnungen zur Weiterentwicklung der EML ASR verwendet wurde, musste die Zustimmung der Vortragenden eingeholt werden. Das war notwendig, da die beim Training gewonnenen statistischen Parameter in der Spracherkennungssoftware gespeichert wurden. Im Zuge dessen wurden jedoch keine personenbezogenen Daten gespeichert und sind somit nicht mehr auf die jeweilige Person zurückzufolgern. Das Einholen der Zustimmung erfolgte zu Semesterbeginn gemeinsam mit der Anfrage zur Durchführung der Aufzeichnungen selbst. Der genaue Ablauf ist in Abschnitt 5.3.1 auf Seite 59 beschrieben.

#### Weiterentwicklung mit Hilfe der gewonnenen Trainingsdaten

Um die Erkennungsraten der EML ASR für den österreichischen Dialekt zu verbessern, stand bis zum Ende des praktischen Teils der Diplomarbeit nur die Möglichkeit des Trainings des Sprachmodells zur Verfügung. Wie im ersten Teil dieses Kapitels im Abschnitt 3.1.2 auf Seite 21 beschrieben, handelt es sich dabei um das im Spracherkennungssystem verwendeten Vokabular, sowie der Wahrscheinlichkeit für das Auftreten einzelner Wörter und Wortfolgen. Der Zugriff und somit das Training des akustischen Modells durch externe Partner war bis dahin

seitens des EMLs noch nicht implementiert.

Für das Training des Sprachmodells wurden die im Zuge der Lehrveranstaltungsaufnahmen aufgezeichneten gesprochenen Vorträge aus den Videodateien extrahiert. Sie wurden unkomprimiert im Wave-Format gespeichert und mittels des Webinterfaces in die EML ASR zum Training eingespielt. Weiters wurden alle verfügbaren, geschriebenen Materialien zum behandelten Themengebiet der Kurse, wie zum Beispiel die Folienpräsentationen, bereitgestellte Skripten, thematisch relevante Quellen im Internet und selbsterstellte Wortlisten in das Spracherkennungssystem hinzugefügt. Dadurch wurden fehlende Vokabel in das Wörterbuch aufgenommen.

Um das Sprachmodell der EML ASR für die Verwendung im Zuge des GESTU Projekts zu trainieren, muss ein möglichst breites Spektrum an Sprechern/-innen und wissenschaftlichen Fachgebieten sowie eine möglichst große Anzahl an Aufnahmen selbst bereit gestellt werden. Weiters ist es im Speziellen für die ersten Audioaufnahmen und Trainingsdaten wichtig, dass für diese zusätzlich 1:1 Transkripte bereitgestellt werden. Sie können entweder händisch oder mithilfe der EML ASR erstellt werden. Da die automatisch erstellten Transkripte jedoch noch sehr viele Fehler enthalten, müssen sie manuell korrigiert werden, um fehlerfreie Texte für das Training der ASR zu erhalten. Beide Erstellungsmethoden sind jedoch sehr aufwändig. Für die manuelle offline Untertitelerzeugung mit herkömmlichen Methoden gibt Romero-Fresco einen durchschnittlichen Erstellungsaufwand von 10:1 an: 10 Minuten Arbeitszeit für eine Minute der zu untertitelnden Aufnahme (siehe [RF11, S: 23]). Aus eigenen Erfahrungen beim manuellen Erzeugen der Transkripte konnte die Rate nur im besten Fall erzielt werden.

Durch das Training mit den bereitgestellten Audioaufnahmen sowie den zusätzlich bereitgestellten Materialien wurden die Wahrscheinlichkeiten für das Auftreten von Wörtern und Wortfolgen im österreichischen Dialekt in das verwendete Sprachmodell miteinbezogen.

### 3.2.2 Ergebnisse

Während des GESTU Projekts wurde eine Vielzahl an universitären Kursen aus den verschiedenen wissenschaftlichen Fachrichtungen aufgezeichnet. Eine detaillierte Beschreibung der Lehrveranstaltungsaufzeichnungen erfolgt in Abschnitt 5.3 auf Seite 59. Insgesamt konnte die Zustimmung von sechs Vortragenden zur Verwendung der Audiomitschnitte ihrer Vorträge zum Training der EML ASR eingeholt werden. Dabei konnte ein breites Spektrum an wissenschaftlichen Fachrichtungen der Sport-, Wirtschafts-, Literaturwissenschaft und der Informatik abgedeckt werden. Somit konnten während des GESTU Projekts über 62 Stunden Tonaufzeichnungen gesammelt werden.

Außerhalb des Modellversuchs wurden an vielen der Wiener Universitäten ebenfalls Lehrveranstaltungen aufgezeichnet. Durch Kooperationen könnte die Anzahl der zum Training verfügbaren Audioaufzeichnungen von Kursen des tertiären Bildungssektors noch vervielfacht werden. Um solch eine Zusammenarbeit anzustreben, wurden mit der Universität für Bodenkultur Wien und der Wirtschaftsuniversität Wien erste Gespräche geführt. Grundsätzlich wurde die Bereitschaft zur Bereitstellung der Aufnahmen zum Training der EML ASR, sofern die jeweiligen Vortragenden damit einverstanden sind, geäußert.

Im Zuge der Diplomarbeit wurde damit begonnen, das Spracherkennungssystem mit einem Teil der gesammelten Vorlesungsaufzeichnungen, selbsterstellten 1:1 Transkripten und allen darüber hinaus verfügbaren Materialien zu trainieren. Dadurch konnten bessere Erkennungsraten gegen-

über dem Sprachmodell für die hochdeutsche Sprache erreicht werden. Bedingt durch den fehlenden Zugang zum akustischen Modell konnten nur die für das Österreichische gebräuchliche Vokabel und Wortfolgen trainiert werden. Ein Training der dialektspezifischen Aussprache von Lauten war dadurch jedoch nicht möglich. Somit war die Weiterentwicklung und Anpassung der EML ASR nur begrenzt möglich. Das zeigte sich durch die erzielten, ungenügenden Worterkennungsraten. Die Auswertung durch das System selbst ergab eine Wortakkuratheit von maximal 50%. Das entspricht einer Verbesserung von durchschnittlich 10 Prozentpunkten gegenüber den Ergebnissen des Standardmodells für die hochdeutsche Sprache.

Laut der Auskunft des European Media Labors sind diese Erkennungsraten, bedingt durch die schlechten Bedingungen beim Einsatz von Spracherkennungssystemen zur Transkription von Vorlesungen durchaus üblich und nicht mit denen von Diktiersystemen vergleichbar<sup>7</sup>. Verglichen mit den erzielten Ergebnisse von sprecherunabhängiger Spracherkennungssoftware für freie englische Sprache, liegen sie im Bereich der von Google entwickelten ASR, die laut Angaben von Romero-Fresco eine WRR von 55,6% erreicht (siehe [RF11, S: 70]). Im Gegensatz zur, vom LLC eingesetzten, „IBM Hosted Transcription Service Engine Controller“ erzielten WRR von ~78,6% sind sie jedoch deutlich schlechter (siehe [RF11, S: 70]).

Die Verwendung von unkorrigierten Transkripten mit einer Wortakkuratheit von 50% ist für den Einsatz zur Untertitelung von Vorlesungen für hörbehinderte Studierende unbrauchbar. Solche Untertitel lassen nur sehr schwer, wenn überhaupt, auf den Inhalt des Vortrages schließen. Exemplarisch ist im Anhang - siehe 9.1 auf Seite 106 - für einen kurzen Ausschnitt einer Lehrveranstaltung das Ergebnis der EML ASR einem händisch erstellten 1:1 Transkript gegenübergestellt. Bedingt durch diese Faktoren und der Konzentration auf die anderen, in dieser Arbeit beschriebenen technischen Hilfsmittel zur Unterstützung hörbehinderter Studierender sowie einer mangelnden Aussicht auf eine rasche Verbesserung der Ergebnisse, wurde schließlich gemeinsam mit Prof. Dr. Wolfgang Zagler die Entscheidung getroffen, die Weiterentwicklung des Spracherkennungsystems im Zuge der Diplomarbeit einzustellen.

### 3.2.3 Empfehlungen

Der Einsatz und die Weiterentwicklung der Spracherkennungssoftware ist zukünftig unter Berücksichtigung einiger Rahmenbedingungen dennoch möglich. Anhand der im vorhergehenden Abschnitt beschriebenen Erkenntnisse und der dort genannten Auskunft des EMLs, ist es absehbar, dass die Verbesserung der Ergebnisse des Spracherkennungssystems für den Einsatz im österreichischen tertiären Bildungssektor nur unter sehr großem Aufwand und der Verwendung einer großen Anzahl an Audioaufnahmen, sowie selbsterstellten Transkripten möglich ist. Ausgehend von dem zuvor beschriebenen Aufwand zur Erstellung von 1:1 Transkripten nimmt deren Erstellung für eine 90 minütige Vorlesung schätzungsweise 900 Minuten oder mehr in Anspruch. Über mögliche Verbesserungen, die durch eine Anpassung des akustischen Modells erzielt werden könnten, kann im Zuge der Diplomarbeit, bedingt durch den fehlenden Zugriff, keine Aussage getroffen werden.

Grundsätzlich ist es denkbar, die durch die EML ASR erstellten Transkripte mit den aktuell erzielten Worterkennungsraten von 50% manuell zu korrigieren, um somit ein akzeptables Er-

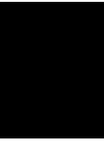
<sup>7</sup> Dabei handelt es sich um eine E-Mail Auskunft vom 11.07.2012 von einem der Entwickler der EML ASR.

gebnis zu erzielen. Das kann natürlich den hörbehinderten Studierenden anschließend auch gemeinsam mit der Vorlesungsaufzeichnung und den Präsentationsfolien in Synote zur Verfügung gestellt werden. Eine Beschreibung dieser speziellen E-Learning Plattform erfolgt in Kapitel 5.2 auf Seite 57. Der benötigte Aufwand für die Durchführung einer manuellen Korrektur lässt sich aus meiner Sicht derzeit nicht abschätzen, denn dahingehend wurden keinerlei Versuche durchgeführt. Ein Einsatz der in Synote neu implementierten Korrektur durch die Studierenden (siehe [WL12]) ist jedoch aus meiner Sicht auf keinen Fall ausreichend, da die durch die EML ASR erzielten Ergebnisse dafür nicht brauchbar sind.

Besteht grundsätzlich Interesse, offline Untertitel von Lehrveranstaltungen für hörbehinderte Studierende zu erstellen, existieren kostengünstige Alternativen. Für den Zweck müsste die zuvor genannte Erstellung eines Transkriptes durch die EML ASR und einer anschließenden manuellen Korrektur auf die grundsätzliche Durchführbarkeit hinsichtlich der dafür benötigten Zeit evaluiert werden. Das muss besonders im Hinblick auf den möglichen Einsatz der Respeaking Technik erfolgen. Laut Romero-Fresco erreicht die Technik bei der Erstellung von Offline-Untertiteln Geschwindigkeiten von 7:1. Das bedeutet zur Erstellung von Untertiteln für eine Minute des Ausgangsmaterials werden sieben Minuten Arbeitszeit benötigt (siehe [RF11, S: 23]). Weiters ist die von Hattinger in seiner zweiten Diplomarbeit [Hat13] erarbeitete Respeaking Ausbildung interessant. Er zeigt die grundsätzliche Durchführbarkeit eines kurzen, maximal sechs Monate dauernden Trainingslehrganges, nach dem es den Teilnehmern/-innen möglich ist, offline Untertitel in sehr guter Qualität zu erstellen.

Auf lange Sicht ist dennoch die Entwicklung auf dem Gebiet der Spracherkennungssysteme zu beobachten. Sofern akzeptable Ergebnisse erzielt werden, ist besonders deren automatischer Einsatz interessant. Wesentlich ist dabei die schnelle, kostengünstige und einfache Erstellung von Transkripten und deren Einbindung in E-Learning Plattformen, wie zum Beispiel Synote. Dadurch lässt sich nicht nur der Nutzen von Aufzeichnungen durch das Zusammenführen mehrerer Informationsquellen, sondern auch die oft geforderte bessere Zugänglichkeit und Suche in Aufzeichnungen erreichen (siehe [Her11, S: 38 - 39], [WWM<sup>+</sup>09]). Somit können nicht nur bestehende Barrieren für hörbehinderte Studenten/-innen verringert werden, sondern auch ein Mehrwert für alle Studierenden geschaffen werden.

KAPITEL 4



**Respeaking**

Im ersten Abschnitt des Kapitels erfolgt eine kurze Einführung in die Respeaking Technik. In den vorhergehenden Kapiteln 2 und 3 wurden Untertitel und Spracherkennungssoftware beschrieben. Beide Themengebiete stellen Grundlagen für die Untertitelerzeugungsmethode dar. Im zweiten Teil wird deren Einsatz zur Echtzeituntertitelung in Lehrveranstaltungen im Zuge des GESTU Projekts und die daraus gewonnenen Erkenntnisse, aber auch Empfehlungen beschrieben. Eine ausführliche Beschreibung des Respeakings ist in der Diplomarbeit von Hattinger [Hat13], die ebenfalls im Zuge des Modellversuchs geschrieben wurde, sowie im Buch von Romero-Fresco [RF11] enthalten.

## 4.1 Geschichte des Respeakings

Zu Beginn des Kapitels 3 wurde festgehalten, dass zur Zeit kein sprecherunabhängiges Spracherkennungssystem existiert, mit dem in Echtzeit aber auch offline vollautomatisch Untertitel für freie Sprache in ausreichend guter Qualität erstellt werden können. Durch den Einsatz unter wesentlich günstigeren Rahmenbedingungen erreichen Diktiersysteme bereits sehr gute Worterkennungsraten bei geringen Verzögerungszeiten. Als Voraussetzung dazu müssen sie jedoch auf eine/n Sprecher/-in trainiert werden und erfordern die Aussprache von Satzzeichen (siehe Kapitel 3.1.1 ab Seite 19). Um solche guten Ergebnisse auch für beliebige Personen und freie Sprache verwenden zu können, wurde um das Jahr 2000 die sogenannte Respeaking Technik entwickelt und vermehrt eingesetzt.

Bedingt durch das gesteigerte Bewusstsein für behinderte Personen und den daraus entstandenen gesetzlichen Vorgaben müssen nationale Fernsehstationen eine gewisse Anzahl ihrer Sendungen mit Untertiteln versehen (siehe Kapitel 2.1 ab Seite 8). Um den großen Bedarf an Echtzeit-Untertiteln decken zu können, waren viele TV Stationen auf der Suche nach brauchbaren Erzeugungsmethoden und haben die verschiedensten Techniken erprobt. Sie waren jedoch entweder nicht leistungsfähig genug, zu teuer oder es stand kaum ausgebildetes Personal zur Verfügung. Vor allem im europäischen Raum setzte sich das Respeaking zur Erzeugung von Live- als auch Offline-Untertiteln durch. Es vereint sehr hohe Genauigkeit mit einer hohen Geschwindigkeit bei vergleichsweise geringen Kosten. Weiters ist die Ausbildungszeit von wenigen Monaten vergleichsweise niedrig (siehe [Hat11, S: 70 - 77], [Hat13, Kapitel 2.1], [RF11, S: 6 - 43]). Ein Überblick über die verschiedenen Methoden zur Untertitelerzeugung ist in Kapitel 2.3 ab Seite 11 beschrieben sowie in Tabelle 2.1 zusammengefasst.

## 4.2 Definition „Respeaking“

Die Respeaking Technik ist noch relativ jung, daher musste erst ein Name gefunden und geprägt werden. Der Name „Respeaking“ ist bei weitem nicht der einzige unter dem die Technik bekannt ist. In diversen Publikationen aber auch in verschiedenen Ländern wurden dafür unterschiedliche Bezeichnungen genannt. So wurde im Englischen unter anderem „speech-based live subtitling“, „speech recognition-based subtitling“, „real-time subtitling via speech recognition“, „speech captioning“ oder „voice-writing“ verwendet. Neben den Bezeichnungen und noch vielen anderen setzte sich im englischen und deutschen Sprachraum der Ausdruck „Respeaking“ durch (siehe [RF11, S: 2 - 5]). Nachdem nun ein Name gefunden wurde, muss noch erklärt und

definiert werden, worum es sich beim Respeaking überhaupt handelt. Romero-Fresco definiert den Ausdruck folgendermaßen:

„A technique in which a respeaker listens to the original sound of a live programme or event and respeaks it, including punctuation marks and some specific features for the deaf and hard of hearing audience, to a speech recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay“ [RF11, S: 1].

Neben dem Satz, der die Erzeugungsmethode bereits gut beschreibt, führt Romero-Fresco noch die wichtigsten Merkmale weiter aus, um den Begriff noch besser zu erklären:

- **Live**  
Das Respeaking wurde zuerst für die Echtzeituntertitelerstellung entwickelt und eingesetzt. Aufgrund der guten Ergebnisse, hohen Geschwindigkeit und geringen Kosten wird es heutzutage jedoch auch zur Erzeugung von Offline-Untertiteln eingesetzt (siehe [RF11, S: 1 - 2]).
- **Respeak**  
„To respeak“ kann man mit „etwas nachsprechen oder wiederholen“ ins Deutsche übersetzen. Bei der Erzeugungsmethode handelt es sich jedoch nicht um ein reines wörtliches Wiedergeben des Gehörten. So muss der/die Respeaker/-in Satzzeichen aussprechen und manchmal Informationen für hörbehinderte Personen hinzufügen. Damit sie mit der zum Teil sehr hohen Sprechgeschwindigkeit mithalten können, müssen sie oftmals kürzen, umformulieren und zusammenfassen, ohne jedoch die Information zu verfälschen. Weiters muss die respeakende Person die Fehler der Spracherkennungssoftware korrigieren, bevor die Untertitel gesendet werden. In den meisten Fällen erfolgt dabei keine Übersetzung. Manchmal wird jedoch auch gleichzeitig in eine andere Sprache übersetzt. Somit vereint das Aufgabenfeld des Respeakings Tätigkeiten der klassischen (offline) Untertitelung und des Simultandolmetschens (siehe [RF11, S: 1 - 2, 45 - 54]).
- **Some specific features for the deaf and hard of hearing audience**  
Untertitel werden sehr oft für hörbehinderte Personen erstellt. Aus dem Grund werden bei deren Erzeugung noch spezielle Hinweise hinzugefügt. Es kann sich um die Kennzeichnung des/r Sprechers/-in oder von sonstigen akustischen Informationen wie zum Beispiel dem Vorkommen von Musik oder Geräuschen handeln (siehe [RF11, S: 1 - 2, 17]).
- **Speech recognition software**  
Beim Respeaking wird eine Spracherkennungssoftware zur Erzeugung der Untertitel verwendet. Man kann sie als alternative Eingabemethode betrachten. Ihre Verwendung bringt auch bestimmte Eigenheiten mit sich, auf die geachtet werden muss. So kann zum Beispiel die Vermeidung gewisser Formulierungen und Wörter zu besseren Erkennungsraten führen. Daher muss der/die Respeaker/-in allgemein über Spracherkennungssysteme Bescheid wissen aber besonders über die Stärken und Schwächen des eingesetzten Modells. Weiters muss er/sie sich und die ASR auf den kommenden Einsatz vorbereiten. Nur dadurch können sehr gute Ergebnisse erzielt werden (siehe [RF11, S: 1 - 2, 45 - 73]).

- **Programme & Subtitles**

Respeaking wird inzwischen nicht mehr nur zur Erzeugung von Untertiteln für Fernsehprogramme verwendet. Es findet inzwischen auch Verwendung bei Konferenzen, an Bildungseinrichtungen, in Museen u.v.m. . Bei diesen Anwendungsfeldern werden die Untertitel ebenfalls eingesetzt, um Inhalte hörbehinderten oder fremdsprachigen Personen zugänglich zu machen (siehe [RF11, S: 1 - 2, 6 - 11, 22 - 40, 138 - 146]).

- **Minimum Delay**

Um den Zusammenhang der akustischen Information mit den anderen Informationsquellen zu gewährleisten, ist eine möglichst geringe Verzögerung bis zur Anzeige der Live-Untertitel wesentlich. Sie wird beim Respeaking von mehreren Faktoren, wie der verwendeten Spracherkennungssoftware oder der Art der Fehlerkorrektur, beeinflusst. So spielt zum Beispiel die verwendete Spracherkennungssoftware oder die Art der Fehlerkorrektur eine wichtige Rolle. Bei der Offline-Untertitelung spielt das selbstverständlich keine Rolle (siehe [RF11, S: 1 - 2, 16 - 17, 45 - 73]).

### 4.3 Respeaking im GESTU Projekt

Zur Zeit werden überwiegend Gebärdensprachdolmetscher/-innen eingesetzt, um gehörlosen Studierenden Zugang zu den nahezu ausschließlich lautsprachlich abgehaltenen Lehrveranstaltungen des österreichischen tertiären Bildungssektors zu ermöglichen. Um die Dolmetschleistungen bezahlen zu können, erhalten hörbehinderte Personen vom Staat finanzielle Unterstützung. Sie reicht jedoch gerade aus, um eine Lehrveranstaltung pro Semester zu besuchen (siehe [KS07, S: 352]). In Wien gibt es derzeit nur ungefähr 20 - 30 ÖGS Dolmetscher/-innen, von denen nicht alle Aufträge im tertiären Bildungssektor annehmen (siehe [KS07, S: 349]). Durch das GESTU Projekt bekamen die teilnehmenden gehörlosen Studierenden erstmals für bis zu fünf Lehrveranstaltungen pro Semester Übersetzungsleistungen zur Verfügung gestellt. Aufgrund des sprunghaft gestiegenen Bedarfes und der geringen Anzahl an qualifizierten Dolmetschern/-innen im Großraum Wien, konnten während des Modellversuchs nicht alle Termine besetzt werden. Bei der Erprobung und Evaluierung technischer Hilfsmittel zur Unterstützung hörbehinderter Studierender im Zuge des GESTU Projekts bestand auch großes Interesse an der Findung von technischen Maßnahmen zum Ausgleich der Sprachbarrieren in den Vorlesungen. Sie sollten dabei als Alternative und ergänzend zu den vorhandenen Hilfsmitteln eingesetzt werden. Laut Hattinger gab es zu Beginn des Modellversuches noch keine Möglichkeit zur sprecherunabhängigen, automatischen Erstellung von Echtzeituntertiteln für freie Sprache für den österreichischen Dialekt beziehungsweise die deutsche Sprache (siehe [Hat11, S: 116 - 124]). Er suchte daraufhin nach anderen Möglichkeiten, um dennoch Live-Untertitel in Lehrveranstaltungen des österreichischen tertiären Bildungssektors einzusetzen. Von den, in Kapitel 2.3 auf Seite 11, genannten Methoden wurde jedoch zu Beginn des GESTU Projekts keine an den Universitäten in Österreich eingesetzt. Daher erprobte Hattinger im Zuge seiner ersten Diplomarbeit am 21.09.2010 gemeinsam mit Mitarbeitern des Modellversuchs den Einsatz von Echtzeituntertiteln in einer auf Englisch abgehaltenen Testvorlesung. Sie wurden mithilfe der Respeaking Technik von der

englischen Firma Red Bee Media<sup>1</sup> erstellt. Durch den ersten Versuch konnte die grundsätzliche Machbarkeit erprobt werden. Die Beurteilung durch die teilnehmenden gehörlosen GESTU Mitarbeiterinnen war ebenfalls sehr positiv (siehe [Hat11, S: 93 - 103]).

### 4.3.1 Durchführung

Ausgehend von dem zuvor erwähnten ersten, positiven Versuch kam es zu einer Kooperation mit der deutschen Red Bee Media Tochterfirma Titelbild<sup>2</sup> mit dem Ziel, die Respeaking Technik an den Wiener Universitäten einzusetzen. Ausgehend davon wurde im Sommersemester 2011 probeweise mit dem Einsatz von Echtzeit-Untertiteln zur Überbrückung der sprachlichen Barriere für eine am Modellversuch teilnehmende Studentin in verschiedenen Vorlesungen an der Universität Wien begonnen. Während des Semesters wurde die Technik im alltäglichen Lehrveranstaltungsbetrieb erprobt. Ein Ziel war es dabei festzustellen, welche Vorlesungstypen und Themengebiete dafür besonders gut geeignet sind und ob es eventuell ungeeignete Inhalte gibt. Während dessen wurden die Qualität und die Verzögerung der Untertitel evaluiert. Dabei spielte besonders die Bewertung durch die hörbehinderten Studierenden eine große Rolle, denn es musste für sie möglich sein, den Lehrveranstaltungen mithilfe der Untertitel zu folgen.

Während des Semesters wurden auch unterschiedliche Methoden der Anzeige ausprobiert. Im Normalfall wurden sie durch einen Laptop angezeigt, der sich vor oder neben den Studierenden befand. In einer Vorlesungseinheit wurden die Untertitel neben der Folienpräsentation des Vortrages für alle sichtbar auf eine Leinwand projiziert. Im Anschluss daran wurde das Feedback vom Vortragenden und allen anwesenden Studierenden eingeholt. Die anwesende hörbehinderte Studentin beurteilte diese Methode als der Anzeige auf dem Laptopbildschirm gleichwertig.

Im Zuge des Probesemesters zeigte sich sowohl die Einsatzfähigkeit im alltäglichen Universitätsbetrieb und dass der Besuch von Vorlesungen durch die ausschließliche Verwendung der Untertitel grundsätzlich möglich ist. Ausgehend von den Erfahrungen des Sommersemesters 2011 wurde der Ablauf zur Planung und zum Einsatz der Echtzeituntertitelung durch die Respeaking Technik erstellt. In der zweiten Hälfte des GESTU Projekts - im Wintersemester 2011/2012 und im Sommersemester 2012 erfolgte bereits eine Betreuung des Einsatzes während der Vorlesungen durch Tutoren/-innen des Modellversuches.

### Gesetzliche Lage

Hattinger beschreibt im Kapitel über die rechtliche Grundlage technischer Hilfsmittel und deren Einsatz im tertiären Bildungssektor seiner Diplomarbeit, dass eine Vorlesung laut dem §2 des österreichischen Urheberrechts ein „mündlich vorgetragenes Sprachwerk“ [Hat11, S: 33] ist. Grundsätzlich ist eine Bearbeitung eines solchen urheberrechtlichen Werkes zulässig. Die Veröffentlichung bedarf jedoch der Zustimmung des/r Urhebers/-in.

Die Erstellung von Echtzeituntertiteln ist eine solche Bearbeitung und der Einsatz im Hörsaal oder in E-Learning-Plattformen entspricht einer Veröffentlichung. Daher bedarf ihre Verwendung der Zustimmung der vortragenden Person (siehe [Hat11, S: 35]).

<sup>1</sup> <http://www.redbeemedia.com> - letzter Zugriff am 06.03.2013.

<sup>2</sup> <http://www.titelbild.de> - letzter Zugriff am 06.03.2013.

### Ablauf zu Semesterbeginn

Die Planung, Koordinierung und Einteilung der Dolmetscher/-innen aber auch der einzelnen technischen Hilfsmittel im Zuge des GESTU Projekts wurden vor dem Semesterbeginn durchgeführt. Die am Projekt teilnehmenden Studierenden gaben die Lehrveranstaltungen bekannt, die sie im kommenden Semester besuchen werden. Im Zuge dessen gaben sie auch an, wo sie welche technischen Maßnahmen verwenden wollen. Daraufhin wurde die grundsätzliche Eignung und Durchführbarkeit des Hilfsmittels im entsprechenden Kurs abgeschätzt und organisiert.

Wie bereits zuvor beschrieben, müssen die Vortragenden der Durchführung der Echtzeituntertitelung ihrer Vorlesung zustimmen. Daher wurden sie von der Behindertenbeauftragten der jeweiligen Universität stellvertretend für das GESTU Projekt um ihr Einverständnis gefragt. Gleichzeitig wurde mit der Firma Titelbild besprochen, welche Termine für sie durchführbar waren.

Erfolgte die Zustimmung der Vortragenden und bestätigte die Respeaking Firma die Termine, übernahmen Tutoren/-innen des GESTU Projekts die Durchführung und Betreuung der Maßnahmen in den Vorlesungen. Die sogenannten eTutoren/-innen wurden nach dem Feststehen aller Lehrveranstaltungen und den eingesetzten technischen Maßnahmen gesucht und eingeteilt. Um die Anzahl der benötigten Tutoren/-innen besser planen zu können, wurde vor jedem Semester der Stundenaufwand für die geplanten technischen Maßnahmen abgeschätzt. Als Grundlage diente dafür die Anzahl der Lehrveranstaltungstermine (**Anzahl**) multipliziert mit der doppelten tatsächlichen Dauer der einzelnen Termine (**Dauer**). Das ist in der nachfolgenden Formel dargestellt. Die Schätzung umfasst die gesamten Tätigkeiten, von der Betreuung der Maßnahme in den Vorlesungen bis zur Abholung und der Rückgabe des Equipments in der GESTU Service-stelle. Die Schätzung deckte sich sehr gut mit dem tatsächlichen Stundenaufwand während des Semesters.

$$\text{Aufwand zur Betreuung einer Lehrveranstaltung} = \text{Anzahl} * \text{Dauer} * 2 \quad (4.1)$$

Die einzelnen Tutoren/-innen wurden vor Semesterbeginn und während der ersten Vorlesungen in das benötigte Equipment und alle notwendigen Arbeitsschritte eingeschult.

Während der ersten Lehrveranstaltungstermine sollten die hörbehinderten Studierenden beurteilen, ob es ihnen durch die Echtzeit-Untertitel möglich ist, dem Inhalt der Vorlesungen zu folgen. War das der Fall, wurde das Hilfsmittel weiter eingesetzt.

### Vorbereitungen auf eine kommende Vorlesung

Zu Beginn des Kapitels wurde bereits beschrieben, dass sich die Respeaker/-innen selbst und die eingesetzte Spracherkennungssoftware auf die kommenden Einsätze vorbereiten müssen. Dazu machen sie sich mit den zu erwartenden Inhalten und Themengebieten vertraut. Dabei fügen sie unter anderem der ASR unbekannte Wörter in das Vokabular ein und erstellen spezielle Makros. Es ist daher vor jeder Lehrveranstaltung wichtig, dass alle Unterlagen der kommenden Vorlesungstermine an die Respeaking Firma zur Vorbereitung übermittelt werden. Das ist die Aufgabe des/der eTutors/-in, der/die durch seine/ihre Anwesenheit in ständigem Kontakt mit den Vortragenden steht und sollte zumindest vier Tage, besser noch eine Woche vor dem Vorlesungstermin geschehen.

### **Ablauf in den Vorlesungen**

Neben der Übermittlung der Unterlagen an die Firma Titelbild vor den eigentlichen Vorlesungsterminen, war der/die eTutor/-in auch für die Durchführung in den einzelnen Lehrveranstaltungen zuständig. Dabei waren sie für die Mitnahme, den Auf- und Abbau und die Betreuung des benötigten technischen Equipments verantwortlich, bestehend aus einem Funkmikrofonset und einem Laptop. Die einzelnen Komponenten und die Anforderungen an sie sind im nächsten Abschnitt detailliert beschrieben. Vor dem Start der Vorlesung wurde die Empfangseinheit des Funkmikrofonsets mit dem Laptop verbunden. Die Sendeeinheit wurde der Vortragenden Person übergeben, die das Mikrofon in der Mitte des Oberkörpers befestigt. Das soll an einem nicht sehr bewegten Kleidungsstück geschehen, zum Beispiel am Sakko. Schals oder Krawatten sind dafür eher ungeeignet. Der Laptop wurde mit dem an den Wiener Universitäten verfügbaren WLAN mit dem Internet verbunden. Mit einem Internetbrowser wurde die Webapplikation zur Anzeige der Untertitel geöffnet. Abschließend wurde mit Skype<sup>3</sup> eine Audioverbindung zur Firma Titelbild hergestellt und eine kurze Überprüfung des Mikrofons und der Tonqualität durchgeführt. Während der laufenden Vorlesung überwachte der/die Tutor/-in das Equipment, achtete auf dessen Funktionsfähigkeit und behob eventuelle technische Probleme.

Durch die zuvor bereits erwähnte Webapplikation wurden die Untertitel mit einer kurzen Verzögerung angezeigt. Es war dabei möglich einige Anzeigeparameter zu verändern:

- **Einstellung der Schriftgröße**
- **Änderung der Text- und Hintergrundfarbe**
- **Auswahl des Zeilenumbruches**  
Hierbei konnten nach jedem Absatz, Satz oder keine zusätzlichen Zeilenumbrüche hinzugefügt werden.
- **Festlegung der Textanzeigegrenze**  
Durch das Festlegen einer horizontalen Grenze war es möglich zu bestimmen, wie viel Text am Monitor zu sehen ist und wie tief dieser bis zum unteren Rand des Monitors angezeigt wird. Das erleichtert den Blickwechsel zwischen Monitor und dem/der Vortragenden oder den Präsentationsfolien.
- **Nummerierung der einzelnen Sätze**  
Mit der Aktivierung dieser Option wurden die Sätze mit einer fortlaufenden Nummer versehen. Somit war es möglich, die Nummer auf den Folien, Notizen oder Bildern zu notieren, um so die entsprechende Stelle in den Untertitel schneller aufzufinden.

Nach dem Ende der Vorlesung wurde das Equipment wieder vom/von der eTutor/-in eingesammelt und in die GESTU Servicestelle zurückgebracht. Die in der Vorlesung erstellten Texte wurden im Anschluss daran von der Firma Titelbild als vollständiges Transkript zur Verfügung gestellt. Es wurde an die hörbehinderten Studierenden und, falls gewünscht, an die Vortragenden übermittelt.

<sup>3</sup> <http://www.skype.com> - letzter Zugriff am 06.03.2013.

Außerdem stellte die Respeaking Firma auch noch die Untertitel inklusive der Zeitcodes bereit. Somit konnten sie in der Synote Plattform (siehe Abschnitt 5.2 auf Seite 57) gemeinsam mit Vorlesungsaufzeichnungen und den Präsentationsfolien synchron abgespielt werden. Der Einsatz und die Versuche mit dieser E-Learning-Plattform sind in Abschnitt 5.3.2 auf Seite 67 beschrieben.

### **Eingesetztes Equipment**

Für die Durchführung der Echtzeituntertitelung durch die Respeaking Technik wird während der Vorlesung lediglich ein Laptop sowie ein Funkmikrofonset benötigt:

- **Computer**

Für die Durchführung des Respeakings in den Lehrveranstaltungen ist ein Computer notwendig. Er dient zur Anzeige der Echtzeituntertitel und zur Übermittlung des gesprochenen Vortrages an die Firma Titelbild. Aufgrund der beiden Aufgaben ist es wichtig, dass der Computer über einen 3,5mm Klinke-Mikrofoneingang und ein WLAN-Modul verfügt. Darüber hinaus müssen die beiden im nächsten Abschnitt genannten Programme mit ausreichender Geschwindigkeit ausführbar sein. Die Anforderungen werden inzwischen von den meisten Laptops erfüllt.

- **Funkmikrofon - Sennheiser ew100 G3, E-Band**

Zum Erreichen einer möglichst geringen Verzögerung und einer hohen Qualität der Untertitel, ist beim Respeaking eine sehr gute Tonqualität notwendig. Daher wurde ein externes Mikrofon eingesetzt, das an der Kleidung der Vortragenden Person befestigt wurde. Es darf jedoch nicht zu groß oder zu schwer sein, damit die Vortragenden nicht in ihrer Bewegungsfreiheit eingeschränkt werden. Sowohl der Sender als auch der Empfänger mussten, bedingt durch den häufigen Transport, möglichst kompakt sein. Weiters mussten sie auch durch Batterien oder Akkus mit Energie versorgt werden können, da es in manchen Hörsälen nur sehr wenige Stromsteckdosen gibt. Dadurch wurde die Auswahl der käuflichen Funkmikrofonsets stark eingeschränkt.

Das gewählte „Sennheiser ew100 G3“ Set erfüllt die zuvor beschriebenen Anforderungen. Im ersten Jahr des Modellversuchs konnte immer wieder ein solches Set von der Universität Wien ausgeliehen und eingesetzt werden. Dabei zeigte sich nicht nur die einfache Handhabung sondern auch die gute Verarbeitungs- und Tonqualität.

Sennheiser bietet für dieses Set verschiedene Modelle an, die in unterschiedlichen Frequenzbereichen arbeiten. Da in Österreich unter anderem der Funkfrequenzbereich von 822 - 832 MHz zur störungsfreien Verwendung von Funkmikrofonen ohne vorherige Anmeldung zur Verfügung steht, wurde beim Kauf das E-Band Modell gewählt. Damit lässt sich eine Frequenz im Bereich von 823 - 865 MHz wählen (siehe [Bun11]).

### Eingesetzte Software

Bei der Durchführung des Respeakings wurden die nachfolgende Software eingesetzt.

- **Internet Browser**

Wie bereits beschrieben, wird von der Firma Titelbild eine eigene Webapplikation zur Anzeige der Untertitel bereitgestellt. Zu deren Verwendung ist ein Internet Browser notwendig. Titelbild empfiehlt dazu den Einsatz von Google Chrome<sup>4</sup>.

- **Skype**

Der gesprochene Vortrag wird durch die Internettelefoniesoftware Skype<sup>5</sup> an die Firma Titelbild übertragen. Deren Einsatz wurde während der ersten Versuche mit Red Bee Media und Titelbild getestet. Es konnte festgestellt werden, dass die Tonqualität und Verzögerung für das Respeaking ausreicht. Skype ist weiters für eine Vielzahl von Betriebssystemen verfügbar, kostenlos einsetzbar und einfach zu bedienen.

#### 4.3.2 Erkenntnisse

Die Echtzeituntertitelung durch die Respeaking Technik wurde im GESTU Projekt mit großem Erfolg eingesetzt. Durch verschiedene Versuche zeigte sich die grundsätzliche Einsetzbarkeit im Studienalltag an den Wiener Universitäten. Die Maßnahme wurde von den hörbehinderten Studierenden angenommen und es war ihnen möglich, dem Inhalt der Lehrveranstaltungen ohne den Einsatz von ÖGS Dolmetschern/-innen zu folgen. Die Erkenntnisse aus der Befragung der Studierenden sind im letzten Teil des Abschnitts zusammengefasst.

Während der ersten Projekthälfte wurde der Ablauf zur reibungslosen Organisation zu Semesterbeginn und zur Durchführung während der Vorlesungen erstellt. Bei der Echtzeituntertitelung durch das Respeaking ist es wesentlich, auf eine gute Aufklärung aller Beteiligten zu achten. Nur dann ist es möglich, dieses Hilfsmittel zu deren Zufriedenheit einzusetzen. Das ist zwar bei allen technischen Maßnahmen der Fall, trifft jedoch auf diese besonders zu. Das beginnt mit den hörbehinderten Studierenden. Bei den Untertiteln existieren oftmals falsche Vorstellungen und überzogene Erwartungen bezüglich deren Qualität und Verzögerung, weil sie sehr oft nur vom Fernsehen oder Filmen bekannt sind. Dabei handelt es sich jedoch meist um vorbereitete, fehlerfreie und perfekt synchronisierte Texte.

Doch auch die vortragende Person muss umfassend über das technische Hilfsmittel und den Grund des Einsatzes informiert werden. Wie bereits zuvor beschrieben, ist das wichtig, da die Vortragenden der Verwendung in der Vorlesung zustimmen müssen. Es kam dabei oft zu Ablehnungen, da die Echtzeituntertitelung großteils unbekannt ist und oft mit einer Aufzeichnung verwechselt wurde. Weiters ist auch die Zusammenarbeit während des Semesters von Bedeutung. Wie bereits in Abschnitt „Vorbereitungen auf eine kommende Vorlesung“ auf Seite 34 beschrieben ist die Übermittlung der Unterlagen zur Vorbereitung auf die kommende Einheit sehr wichtig. Das funktionierte während des GESTU Projekts großteils sehr gut. Dennoch war die Kooperation in manchen Lehrveranstaltungen nicht reibungslos möglich. Das zeigte deutlich, wie wichtig das Entgegenkommen der vortragenden Person ist und, dass eine fehlende

<sup>4</sup> <http://www.google.de/intl/de/chrome/browser> - letzter Zugriff am 06.03.2013.

<sup>5</sup> <http://www.skype.com> - letzter Zugriff am 06.03.2013.

Vorbereitung die Qualität der Untertitel negativ beeinflusst.

Die Zusammenarbeit mit der beauftragten Respeaking Firma Titelbild funktionierte sehr gut. Die Planung und Durchführung der Einsätze an den Universitäten erforderte sehr viel Flexibilität. Besonders bis zur endgültigen Fixierung aller Lehrveranstaltungen, in denen das technische Hilfsmittel eingesetzt werden soll, kann es sehr oft zu erheblichen Änderungen sowohl bei der Anzahl als auch beim Zeitpunkt der Termine selbst kommen. Sehr oft bereitete die späte Zustimmung von Vortragenden große Probleme.

Insgesamt wurde die Echtzeituntertitelung durch die Respeaking Technik während des Modellversuchs in verschiedenen Vorlesungen an den Wiener Universitäten erprobt. Es zeigte sich vor allem, dass das Respeaking nicht für alle Lehrveranstaltungstypen und Inhalte geeignet ist. Die besten Bedingungen für den Einsatz existieren bei Vorträgen mit wenig Interaktion und Multimediainhalten wie zum Beispiel Bildern, Audio- oder Videoeinspielungen. Beim Einsatz dieses Hilfsmittels wird die akustische Information von einem Mikrofon, das der/die Vortragende bei sich trägt, aufgezeichnet und an die Respeaker/-innen übermittelt. Manche Wortmeldungen werden vom Mikrofon zu leise oder gar nicht übertragen und können somit auch nicht in den Untertiteln erfasst werden. Dasselbe gilt für Audio- und Videoeinspielungen. Daher ist es wichtig, dass Publikumsmeldungen von der Vortragenden Person wiederholt werden. Bei schnell wechselnden Bildern ist die Verzögerung bis zur Anzeige der Untertitel problematisch. Dadurch kann es vorkommen, dass die aktuell angezeigten Sätze zu dem zuvor gezeigten Bild gehören. Das ist jedoch nur bei schnellen Bildwechseln der Fall - also wenn diese weniger als 20 Sekunden gezeigt werden.

Neben der Gestaltung des Vortrages selbst wirken sich auch Pausen sowie ein langsames oder normales Sprechtempo positiv auf die Qualität der Untertitel aus. Im Gegensatz dazu kann sich ein abschweifender Erzählstil durch das möglicherweise der Spracherkennungssoftware nicht trainierte Vokabular negativ auswirken. Durch die Versuche zeigte sich auch, dass das Respeaking für manche Vorlesungsinhalte ungeeignet ist. In dem Zusammenhang sind besonders mathematische Inhalte zu nennen. Sie sind sehr schwer zu diktieren und selbst kleinste Fehler in Formeln ändern deren Bedeutung. Abwechselnde Sprecher/-innen, viel Interaktion oder in der Vorlesung gezeigte Versuche sind beim Einsatz dieser Maßnahme ebenfalls problematisch.

Die zuvor beschriebenen Rahmenbedingungen und Erfahrungen haben großen Einfluss auf die Qualität und Verzögerung der Untertitel. Im schlimmsten Fall können sie den Einsatz und die Brauchbarkeit dieses Hilfsmittels gefährden. Es handelt sich dabei jedoch um Faktoren, die größtenteils vor dem Semesterbeginn bei der Organisation berücksichtigt werden können. Manche der zuvor genannten negativen Einflüsse können auch während des Semesters in Absprache mit den Vortragenden verbessert werden.

Während der Versuche und der Verwendung der Echtzeituntertitel kam es dabei meistens zu einer Verzögerung von durchschnittlich 13 - 18 Sekunden, in Extremfällen bis zu 30 Sekunden bis zur Anzeige der Untertitel. Das wurde anhand von Stichproben während ausgewählter Versuche gemessen. Die Qualität der Untertitel selbst wurde ebenfalls anhand subjektiver Eindrücke stichprobenartig während des Einsatzes aber auch durch die NER Methode, die in Abschnitt 2.4 auf Seite 14 beschrieben ist, festgestellt. Somit konnte Hattinger für eine ausgewählte Testvorlesung einen NER Wert für die Untertitel der Firma Titelbild von ~95% errechnen (siehe [Hat13, Kapitel 4]). Der Wert liegt unter dem von Romero-Fresco genannten Grenzwert von 98% für

Untertitel in guter Qualität (siehe [RFM13]). Dennoch sind diesbezüglich einige erschwerende Faktoren bei der gewählten Vorlesung selbst zu berücksichtigen. Vor allem war das Sprechtempo des Vortragenden mitunter sehr hoch (siehe [Hat13, Kapitel 4]) und die vorab zur Verfügung gestellten Unterlagen deckten oftmals nicht den gesamten in der Vorlesung behandelten Inhalt und Stoff ab. Im Anhang 9.1 auf Seite 106 ist ein Abschnitt aus den Respeaking-Untertiteln einer manuell erstellten 1:1 Transkription gegenübergestellt. Im Allgemeinen zeigte sich bei den einzelnen Versuchen, dass die Vereinfachungen und Kürzungen bei schlechter Tonqualität, abgebrochenen oder unvollständigen Sätzen, aber auch bei zu schnellem Vortrag, mitunter stark zunehmen.

### **Bewertung durch die hörbehinderten Studierenden**

Bei allen technischen Hilfsmitteln, die im Zuge des GESTU Projekts erprobt und eingesetzt wurden, stellt die Rückmeldung der am Projekt teilnehmenden Studierenden einen maßgeblichen Anteil zu deren Beurteilung dar. Es wurden daher die beiden Studentinnen, welche die Echtzeituntertitel in Vorlesungen erprobten, zu ihren Erfahrungen und ihrer Meinung zum Einsatz dieser technischen Maßnahme befragt. Die Echtzeituntertitel wurden in insgesamt fünf Lehrveranstaltungen eingesetzt. Die Befragung wurde für jede Vorlesung getrennt durchgeführt und aufgezeichnet.

Eine der beiden befragten Studentinnen bezeichnet sich selbst als schwerhörig, die zweite als gehörlos. Trotzdem folgten beide normal dem gesprochenen Vortrag in Lehrveranstaltungen mithilfe von ÖGS Dolmetschern/-innen. Ohne die Dolmetscher/-innen können sie den Inhalten eines Kurses laut eigenen Angaben nur sehr schlecht folgen. Ihre Lesekompetenz beurteilen beide als gut bis sehr gut und gaben an, dass diese ihrer Meinung nach keinen Einfluss auf die Verständlichkeit der Untertitel hatte.

Die Untertitel, die durch die Respeaking Technik erstellt wurden, wurden allgemein positiv beurteilt:

- In allen 5 Lehrveranstaltungen beurteilten die Studentinnen die Untertitel als „brauchbar“.
- Der gesprochene Inhalt von 3 Vorlesungen wurde durch die Untertitel „gut“ und in 2 „ausreichend“ zugänglich gemacht.

Eine der beiden Studierenden konnte nach eigenen Angaben dem Inhalt von den drei von ihr besuchten Kursen mithilfe der Untertitel im Vergleich zu ihren regulär verwendeten Hilfsmitteln (zum Beispiel ÖGS, Lippenlesen,...) besser folgen. Die zweite Studentin gab bei der Frage an, dass sie in den beiden von ihr besuchten Vorlesungen, in denen die Untertitel eingesetzt wurden, den Inhalten gleich gut folgen konnte.

Die beiden Studierenden verwendeten die Echtzeituntertitel immer gemeinsam mit anderen Kommunikationskanälen. Die schwerhörige Studentin gab dabei an, dass sie in den drei von ihr besuchten Vorlesungen zusätzlich zu den Untertiteln noch dem lautsprachlichen Vortrag folgte und von den Lippen des Professors ablas. Je nach Vortragsstil benutzte sie die Untertitel unterschiedlich häufig. Dabei schätze sie die anteilige Verwendung der Untertitel mit 20%, 60% und 85%.

Die zweite Studentin gab hierbei an, dass sie in beiden Vorlesungen zu 75% dem ÖGS Tutor folgte und die Untertitel zu 25% verwendete. Sie hatte laut eigenen Angaben Probleme bei der Verwendung der Untertitel während der Vorlesung und des Transkripts danach. Sie beurteilte die Untertitel im Vergleich zur anderen Studierenden schlechter, gab jedoch bei der Befragung folgendes an:

„Meiner Meinung nach ist die ÖGS besser für mich, aber wenn keine Dolmetscher/-innen zur Verfügung stehen, ist Respeaking eine gute Alternative.“

Im Zuge der Befragungen gaben die beiden Studentinnen folgende Bewertung der Qualität der Untertitel ab:

- Insgesamt wurde die Qualität der Untertitel dreimal mit „gut“ und zweimal „mittel“ beurteilt.
- Die sprachliche Qualität wurde dreimal mit „gut“, einmal mit „mittel“ und einmal mit „schlecht“ bewertet.
- Die Fehlerrate wurde zweimal mit „gering“, zweimal mit „mittel“<sup>6</sup> und einmal mit „hoch“ beurteilt.
- Auf die Frage inwiefern die Fehler die Verständlichkeit beeinflussen wurde zweimal „wenig“, dreimal „mittel“ und einmal „stark“<sup>7</sup> angegeben.
- Insgesamt wurden die Fehler als in einer Vorlesung „gar nicht“, in zwei als „wenig“ und in jeweils einer als „mittelmäßig“ beziehungsweise „stark“<sup>8</sup> störend empfunden.
- Bei der „gar nicht“ und den beiden „wenig“ Nennungen der vorherigen Frage gab die Studentin an, dass sie die Fehler aus dem Kontext selbstständig korrigieren konnte.

Die Verzögerung wurde zweimal als „normal“ und dreimal als „lange“ bewertet. Sie wurde durchschnittlich auf 10 - 15 Sekunden geschätzt. Das deckt sich mit den durchgeführten Messungen. Die schwerhörige Studentin gab an, dass sie die Verzögerung bei gleichzeitigem Lippenlesen als störend empfand.

Als Vorteile beim Einsatz der Echtzeituntertitel nannten die Beiden folgende Punkte:

- Es wird der gesamte Vortrag schriftlich übersetzt.
- Mithilfe der Untertitel ist es möglich, verpasste Inhalte nachzulesen.
- Es ist anhand die Untertitel möglich, dem Vortrag kurz nicht zu folgen und im Anschluss den versäumten Text nachzulesen. Durch die kurzen Pausen ist eine bessere Konzentration möglich.

<sup>6</sup> Eine der beiden Bewertungen wurde durch viele Auslassungen begründet.

<sup>7</sup> Diese Bewertung erfolgte in der Vorlesung, die laut der Studentin eine „schlechte“ Fehlerrate aufwies.

<sup>8</sup> Diese Bewertung erfolgte in der Vorlesung, die laut der Studentin eine „schlechte“ Fehlerrate aufwies.

- Das Transkript erlaubt eine bessere Nachbearbeitung der vergangen Vorlesung sowie eine bessere Vorbereitung auf die kommende Einheit. Von beiden Studentinnen wurde das bei allen fünf Vorlesungen genannt.

Die folgenden Punkte wurden als nachteilig empfunden:

- Die Webapplikation springt bei Erscheinen eines neuen Satzes sofort zur aktuellen Stelle. Folglich ist es nicht oder nur sehr begrenzt möglich, die vorhergehenden, nicht mehr angezeigten Untertitel nachzulesen.
- Auslassungen und nicht übersetzte Passagen sind in den Untertiteln nicht erkennbar.
- Manche Fehler sind anhand des Kontexts nicht erkenntlich oder beim Lesen verwirrend.
- Bei Verbindungsproblemen zur Webapplikation sind die bisherigen Untertitel während der Vorlesung nicht mehr einsehbar.

Abschließend wurden beide Studentinnen noch um ihre Vorschläge zur Verbesserung dieses technischen Hilfsmittels gefragt. Dabei schlug eine der beiden vor, die Transkripte auf einer Internet-Plattform allen hörbehinderten Studierenden zur Verfügung zu stellen. Das wäre für sie auch hilfreich, sofern sie die Texte versehentlich löscht. Die zweite Studentin schlug vor, die vollständigen Transkripte korrigieren und formatieren zu lassen, um sie dadurch besser zum Lernen nutzen zu können.

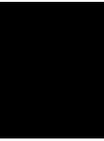
### 4.3.3 Empfehlungen

Durch die Versuche und die Verwendung der Echtzeituntertitel im Zuge des GESTU Projekts und die positive Beurteilung durch die beiden hörbehinderten Studentinnen zeigte sich deren Eignung zum Einsatz im tertiären Bildungssektor. Doch die Untertitelerzeugung in Vorlesungen ist mitunter eines der schwierigsten Anwendungsgebiete. Das begründet sich vor allem aus der wissenschaftlichen Fachsprache, den breiten, vielfältigen und oftmals komplexen Themengebieten und der Unvorhersehbarkeit der Inhalte. Es ist deshalb bei der Planung vor dem Semesterbeginn besonders wichtig, sämtliche Rahmenbedingungen zu überprüfen, alle Beteiligten über das Einsatzgebiet, die Möglichkeiten und Ziele der Technik aufzuklären und somit möglichst gute Bedingungen aber auch angemessene Erwartungen für den Einsatz zu schaffen. Die hörbehinderten Studierenden sollten während der ersten Vorlesungstermine besonders gut betreut werden. Dadurch kann die Eignung der Echtzeituntertitel für die betreffende Person festgestellt werden beziehungsweise können im Bedarfsfall Änderungen vorgenommen werden. Somit kann sichergestellt werden, dass der gesprochene Vortrag für sie durch die Untertitel adäquat zugänglich gemacht wird.

Bei Einhaltung der angeführten Bedingungen stellen die Echtzeituntertitel eine gute Alternative zu Gebärdensprachdolmetschern/-innen dar. Das bietet die Chance und Möglichkeit eventuelle Probleme bei der Rekrutierung von Dolmetschern/-innen auszugleichen. Der große Vorteil des Respeakings besteht jedoch darin, Vorlesungsinhalte unabhängig von den ÖGS Kenntnissen des/der jeweiligen Studierenden zugänglich zu machen.



KAPITEL 5



**Lehrveranstaltungsaufzeichnung**

Im vorliegenden Kapitel werden die im Zuge des GESTU Projekts durchgeführten Lehrveranstaltungsaufzeichnungen, deren Organisation, die Ausführung und die Erkenntnisse behandelt. Bei den Aufnahmen wurde speziell auf die Bedürfnisse der hörbehinderten Studierenden Rücksicht genommen. Zu Beginn werden Vorlesungsaufzeichnungen allgemein behandelt. Während der Literaturrecherche zeigte sich, dass es sehr viele wissenschaftliche Publikationen auf diesem Gebiet gibt. Sie behandeln jedoch nur selten das Thema aus der Sicht von behinderten oder gar hörbehinderten Studierenden. Somit wird im ersten Abschnitt des Kapitels die Lehrveranstaltungsaufzeichnung allgemein behandelt. Es werden die aktuellen Entwicklungen, die Anforderungen, die Durchführung aber auch die Auswirkungen durch den Einsatz erörtert. Abschließend wird dieses Thema aus der Sicht der hörbehinderten Studierenden beschrieben.

## 5.1 Stand der Technik und Forschung

Durch die Entwicklungen auf dem Gebiet der modernen Computer- und Multimediatechnik wird die Aufzeichnung von Lehrveranstaltungen immer einfacher und zugleich populärer. Dabei spielt die Verfügbarkeit billiger und leistungsfähiger Hardwarekomponenten und besonders die weite Verbreitung von Breitband-Internetzugängen eine große Rolle. Zusätzlich besitzen Studierende inzwischen sehr oft einen eigenen Computer oder mobile Geräte wie zum Beispiel Laptops, Tablet-PCs oder Smartphones.

In den Anfängen wurden die einzelnen Kurse noch durch herkömmliche Fernsehkameras auf Kassetten aufgezeichnet, nachträglich digitalisiert, nachbearbeitet und auf Speichermedien zur Verfügung gestellt. In den späten 1990er Jahren wurden verschiedene Versuche zur Vereinfachung und Automatisierung der Lehrveranstaltungsaufzeichnungen durchgeführt. In den letzten zehn Jahren etablierten sich, oftmals aus diesen Projekten, verschiedene Systeme zur automatischen Aufnahme. Deren Ziel ist es, sämtliche Tätigkeiten mit minimalem manuellen Aufwand durchzuführen. Die somit produzierten Aufzeichnungsdokumente werden den Studierenden im Anschluss über das Internet in speziellen Plattformen zur Verfügung gestellt. Durch die Entwicklungen der letzten Jahre wird nicht nur die Durchführung einfacher und effektiver, auch die Möglichkeiten der E-Learningumgebungen werden immer vielfältiger. Es ist inzwischen möglich, Kapitelmarken einzufügen, mehrere Quellen einzubinden und synchron abzuspielen, die Inhalte zu durchsuchen und Notizen einzubinden (siehe [Echb], [FHK08], [Her11, S: 14 - 22], [McC08]).

Das trägt, vor allem bei den Studierenden, zur Popularität von Lehrveranstaltungsaufzeichnungen maßgeblich bei und sehr viele Bildungseinrichtungen auf der ganzen Welt wenden diese Technik in ihren Hörsälen an. Toppin fasst die aktuelle Situation mit folgendem Satz zusammen:

„[...] lecture capture has been gaining momentum, but that momentum is being out-paced by student demand“ [Top11].

Der Begriff der Lehrveranstaltungs- und Vorlesungsaufzeichnung selbst wird für eine breite Palette unterschiedlichster Maßnahmen und Tätigkeiten verwendet. Im klassischen Sinne handelt es sich um die Audio- oder Videoaufzeichnung des Vortrages und der Geschehnisse in einer Lehrveranstaltung. Die Aufzeichnung wird den Studierenden im Anschluss daran zur Verfü-

gung gestellt. Toppin verwendet die folgende Definition für den Begriff der Lehrveranstaltungsaufzeichnung:

„[...] an umbrella term describing any technology that allows instructors to record what happens in their classrooms and make it available digitally“ [Top11].

Die Durchführung der Aufnahmen in der Praxis kann dabei mitunter sehr unterschiedlich ausfallen und hängt von verschiedenen Faktoren ab. Sie kann manuell oder automatisiert erfolgen. Es kann nur der gesprochene Inhalt oder ein Videomitschnitt der Präsentation, des/der Vortragenden oder des Tafelbildes erfolgen. Die Aufzeichnungsdokumente können als Dateien zum Download oder als Stream zur Online-Betrachtung in speziellen E-Learning-Plattformen zur Verfügung stehen. Dort können sie auch mit den verschiedensten Kapitelmarkierungen, Annotationen und Kommentaren versehen werden. Mertens definiert daher den Begriff der Lehrveranstaltungsaufzeichnung detaillierter:

„Der Begriff Vorlesungsaufzeichnung umspannt ein weites Feld unterschiedlichster Endprodukte, die von reinen Audioaufzeichnungen bis hin zu multimedial aufbereiteten Dokumenten mit eingebetteten Videodateien reichen. Auch das Ausgangsmaterial einer Aufzeichnung wie etwa die Art der verwendeten Vortragsmaterialien und die Form der Vorlesung können variieren“ [Mer07, S: 21].

### 5.1.1 Motivation für den Einsatz von Vorlesungsaufzeichnungen

Die Gründe für die Aufzeichnung von Lehrveranstaltungen können sehr vielfältig sein. Meistens ist die Hauptmotivation hinter dem Einsatz das Schaffen besserer Möglichkeiten zum Erlernen und Vertiefen des Stoffes. Es handelt sich jedoch meistens nicht um das einzige Motiv:

- Angebot von zusätzlichen Möglichkeiten zur Vertiefung des Stoffes (siehe [Mer07, S: 52 - 54]).
- Verbesserung der Zugänglichkeit des Inhaltes einer Lehrveranstaltung für behinderte Studierende (siehe [Cha07], [Wil06]).
- Bessere Nutzung der Präsenzzeit (siehe [BA04], [DCL09], [Tal12]).
- Steigerung der positiven Abschlüsse und Senkung der Kursabbrüche (siehe [Top11]).
- Nutzung der ständigen, an den Universitäten stattfindenden Erzeugung von Inhalten (siehe [Her11, S: 1]).

Neben den oben genannten Beweggründen gibt es noch weitere, die von Vortragendem/-r zu Vortragendem/-r variieren.

### 5.1.2 Nutzungsszenarien für Lehrveranstaltungsaufzeichnungen

Wird eine Aufzeichnung durchgeführt, sollte sie gezielt in den Ablauf der Lehrveranstaltung eingebunden werden. Andernfalls besteht die Gefahr, dass die erhofften positiven Effekte ausbleiben und sie sich im schlimmsten Fall sogar nachteilig auswirkt. Secker et al. fassen dies in folgender Forderung zusammen:

„Lecture capture systems must be embedded in a larger methodology to support active learning in the classroom“ [SBG10].

Es kann sich dabei um den Einsatz von begleitenden Tutorien, Fragerunden, Diskussionen, Übungen, Projekten oder Tests handeln (siehe [Mer07, S: 35 - 50]). Dadurch sind die Studierenden gefordert, sich aktiv mit dem Stoff auseinanderzusetzen. Bei der Analyse von Hermann zeigte sich zum Beispiel, dass vor Übungsabgaben und Tests vermehrt auf die Aufzeichnungen zugegriffen wurde (siehe [Her11, S: 43 - 54]). Das wird auch in diversen Befragungen von Studenten/-innen angegeben (siehe dazu Abschnitt „Wahrnehmung, Nutzungsverhalten und Auswirkung“ ab Seite 51).

Durch den Einsatz von Lehrveranstaltungsaufzeichnungen können jedoch auch alternative Unterrichts-Konzepte umgesetzt werden, wie das Konzept des sogenannten „Inverted Classrooms“. Dem liegt ein altbewährter Ansatz zugrunde, der auch in den Rechts- und Wirtschaftswissenschaften eingesetzt wird. Dabei erhalten die Studierenden Materialien zum selbstständigen Erarbeiten eines Themas oder Stoffgebietes. Im Falle des „Inverted Classrooms“ handelt es sich bei den Materialien um Aufzeichnungsdokumente. Sie dienen als Grundlage für Aktivitäten während des nächsten Präsenztermins - zum Beispiel bei Diskussionen oder Übungen. Die Methode hat laut Talbert mehrere Vorzüge. Die Zeit, in der sich der Vortragende und die Studenten/-innen begegnen, lässt sich besser nutzen. Weiters zeigen die Ergebnisse, dass der zu erlernende Stoff besser beherrscht und verstanden wird (siehe [Tal12]). Beim Modell des „Inverted Classrooms“ handelt es sich nur um eine mögliche Ausprägung eines der vier, von Mertens beschriebenen, grundlegenden Nutzungsszenarien (siehe [Mer07, S: 35 - 50]):

- **Aufzeichnungen als ergänzendes Material**

Es handelt sich um die klassische und am weitesten verbreitete Einsatzmöglichkeit. Gleichzeitig ist sie auch die mit dem geringsten Aufwand und Risiko. Die Vorlesung wird aufgezeichnet und die Aufnahmen danach den Studierenden - zumeist online - zur Verfügung gestellt. Sie stellen bei diesem Szenario ein ergänzendes Material dar, sollen jedoch nicht den Präsenztermin der Lehrveranstaltung ersetzen. Deshalb sind selten gravierende Anpassungen des didaktischen Konzepts notwendig. Es ist auch kaum notwendig, die Aufnahmen in größerem Umfang nachzubearbeiten (siehe [Mer07, S: 35 - 39]).

- **Aufzeichnungen als zeitfix präsentierter Vorlesungersatz**

In diesem Nutzungsszenario ersetzt die Vorführung der Aufzeichnung, die von einer Vortragenden Person gehaltene Vorlesung komplett. Somit können einzelne Lehrveranstaltungstermine trotz der Abwesenheit des/der Vortragenden angeboten werden, wenn er/sie verhindert ist. Das kann jedoch auch für den gesamten Kurs erfolgen, wenn zum Beispiel kein ausreichend großer Hörsaal gefunden werden konnte (siehe [Mer07, S: 39 - 43]).

- **Aufzeichnungen als zeitvariabel präsentierter Vorlesungersatz**

Hierbei wird die Aufzeichnung ebenfalls als Ersatz für eine, von einer vortragenden Person gehaltene Vorlesung eingesetzt. Der große Unterschied ist jedoch, dass die Vorführung der Aufnahme nicht zu einem bestimmten, vorgegebenen Zeitpunkt stattfindet. Die Studierenden eignen sich den zu erlernenden Stoff selbstständig durch Ansehen der Aufzeichnungen zeit- und ortsunabhängig an. Mertens empfiehlt die Begleitung des Kurses durch Übungen, Tutorien oder Fragerunden. Da bei dieser Nutzungsart den Aufzeichnungen eine völlig andere und wesentlich zentralere Rolle bei der Wissensvermittlung zukommt, sollten sie auch speziell für diesen Zweck konzipiert und erstellt werden (siehe [Mer07, S: 43 - 47]). Beim zuvor genannten Modell des „Inverted Classrooms“ handelt es sich um eine Variante dieses Szenarios.

- **Vorlesung „on Demand“**

Beim vierten von Mertens genannten Szenario werden Aufzeichnungen in einem Archiv gesammelt und den Studierenden zugänglich gemacht. Sie können auf einzelne Einheiten unabhängig von Lehrveranstaltungen zugreifen und die Inhalte eigenständig erarbeiten. Somit können gezielt einzelne Stoffgebiete erlernt oder nochmals aufgefrischt werden (siehe [Mer07, S: 47 - 50]).

### 5.1.3 Bedürfnisse der Anwender/-innen und Nutzer/-innen

Nachdem nun zu Beginn des Kapitels die Motivation aber auch die verschiedenen Nutzungsszenarien für Lehrveranstaltungsaufzeichnungen behandelt wurden, ist es wichtig zu erfassen, worauf bei der Erstellung der Aufnahmen geachtet werden sollte. Die Bedürfnisse teilen sich dabei grundsätzlich in die Sicht der Studierenden (der Nutzer/-innen der Endprodukte) und in die der Lehrenden (der Anwender/-innen der Technik).

#### Bedürfnisse der Studierenden

Um die Anforderungen der Studierenden an die Lehrveranstaltungsaufzeichnung zu erfassen, wurden verschiedenste Befragungen, Studien und Experimente durchgeführt. Dabei wurden folgende Erkenntnisse gewonnen:

- **Welche Informationskanäle werden aufgezeichnet**

Für die Studierenden ist es vor allem wichtig, was aufgezeichnet wird und nachträglich verfügbar ist. Bei der von Hermann durchgeführten Studie bewerteten die befragten Studenten/-innen die Aufzeichnung des gesprochenen Vortrags, gefolgt von der Bildschirmpräsentation und den möglichen Annotationen darauf als die wichtigsten Elemente. Die Aufnahme der vortragenden Person selbst wird mit nur 10% als unwichtig bewertet (siehe [Her11, S: 33 - 58]). McClure betont in diesem Zusammenhang, dass besonders die Audioqualität von großer Bedeutung für die Verwendung der Aufzeichnungen ist (siehe [McC08]).

- **Gute Suche und Navigationsmöglichkeiten**

Neben der Frage, welche Informationsquellen aufgezeichnet werden, spielt vor allem der Zugriff auf einzelne Stellen und die Navigation in den Aufzeichnungsdokumenten eine große Rolle:

„Da Aufzeichnungsdokumente in der Regel nicht wie Spielfilme linear von Anfang bis Ende betrachtet werden, kommt den Navigations- und Interaktionsfunktionen eine besondere Bedeutung zu“ [Her11, S: 38].

Laut Hermann bewerteten 71% der befragten Studierenden die Navigation durch Kapitelmarken, 85% die Navigation durch die Folien, 70% das visuelle Scrollen und 64% die Suche nach Textstellen als wichtig oder sehr wichtig (siehe [Her11, S: 38 - 39]).

- **Einfügen eigener Anmerkungen**

Für Studierende ist es weiters auch wichtig, die einzelnen Aufzeichnungsdokumente mit eigenen Anmerkungen wie zum Beispiel Kommentaren, Notizen oder Bookmarks versehen zu können. Es sollte möglich sein, sie zu durchsuchen und die Aufnahmen an der eingefügten Stelle fortzusetzen. Zusätzlich bieten Anmerkungen - sofern sie mit anderen geteilt werden können - die Möglichkeit, von den Lehrenden über schwer oder unverständliche Passagen Aufschluss zu erhalten und die Gelegenheit, Aufklärung oder zusätzliches Material bereitzustellen (siehe [Her11, S: 38 - 39], [WWM<sup>+</sup>09]).

- **Flexible Wiedergabemöglichkeit**

In der Literatur wird häufig betont, dass die Aufzeichnungen und die E-Learning-Plattformen, worin sie eingebunden werden, einfach zu bedienen sein sollen und auf die eigenen Bedürfnisse angepasst werden können. Dabei kann es sich zum Beispiel um die flexible Anordnung der einzelnen Medien handeln. Weiters ist wesentlich, dass die Wiedergabesoftware für die verschiedenen Betriebssysteme sowie Endgeräte verfügbar und verwendbar ist (siehe [MPGH12]).

- **Barrierefreiheit**

Zuletzt ist auch der barrierefreie Zugang zu den angebotenen Vorlesungsaufzeichnungen aber auch der damit verbundenen Wiedergabesoftware sowie zu E-Learning-Plattformen besonders für behinderte Studierende von Bedeutung (siehe [MPGH12]).

### Bedürfnisse der Lehrenden

Die Bedürfnisse der Lehrenden unterscheiden sich von denen der Studierenden. Sie wollen in der Regel ihre Vorlesung möglichst unkompliziert, ohne Störung oder Änderungen aufzeichnen. Das spiegelt sich auch in ihren Anforderungen an die Aufzeichnungen wieder:

- **Einfache Erstellung der Aufzeichnung**

Die meisten Lehrenden wollen ihre Vorlesung möglichst einfach, unkompliziert und so unaufdringlich wie möglich aufzeichnen oder aufzeichnen lassen. Das muss sich auch in den eingesetzten Systemen zur Aufzeichnung beziehungsweise den Abläufen widerspiegeln (siehe [MPGH12]).

- **Mobile Lösungen**

Es gibt eine Vielzahl an Systemen, mit denen Lehrveranstaltungsaufzeichnungen ohne großen Aufwand durchgeführt werden können. Für deren Einsatz müssen meistens die einzelnen Hörsäle damit ausgestattet werden. Um auch Vorlesungen in anderen Räumen einfach und komfortabel aufzeichnen zu können, ist es sinnvoll, auch mobile Systeme zur Verfügung zu stellen (siehe [MPGH12]).

- **Kompatibilität für die Mediennutzung**

Bei einer Aufnahme sollten möglichst alle, vom/von der Lehrenden im Zuge seines/ihrer Vortrages verwendeten Medien genutzt und aufgezeichnet werden. Dabei kann es sich um Multimediainhalte, Bildschirmpräsentationen aber auch das Tafelbild, Folienpräsentationen oder ein vorgeführtes Experiment handeln (siehe [MPGH12]).

- **Nachbearbeitung**

Manchmal ist es notwendig bestimmte Passagen aus der gerade aufgezeichneten Einheit zu entfernen. Es kann sich um Pausen, Störungen aber auch Urheberrechtlich geschütztes Material oder persönliche Details handeln. Für diese Fälle ist es wichtig, dass eine einfache und gute Möglichkeit zur Nachbearbeitung der Aufnahmen existiert (siehe [MPGH12]).

- **Einfügen von Anmerkungen**

Für manche Lehrenden ist es auch wesentlich, Anmerkungen wie zum Beispiel Kommentare, ergänzendes Material oder Kapitelmarken nachträglich in die Aufzeichnungen einzufügen. Das sollte im eingesetzten System einfach zu handhaben sein (siehe [MPGH12]).

#### 5.1.4 Arten der Aufzeichnung

In den Anfängen der Lehrveranstaltungsaufzeichnung wurde vor allem mit dem Equipment der klassischen Film- und Fernsehproduktion gearbeitet. Die Geräte mussten zum Teil von speziellem Personal betreut werden. Es stellte sich jedoch schnell heraus, dass diese Art der Aufnahme von Vorlesungen nicht optimal ist. Vor allem störten die eingesetzten und damals noch sehr auffälligen Geräte häufig den Vortragenden. Auch die Betreuung und Nachbearbeitung war sehr kostspielig und für eine Aufzeichnung vieler Kurse ungeeignet. Mitte der 1990er Jahre starteten daher Bemühungen und Versuche, alle notwendigen Schritte so weit es geht zu automatisieren. Heute sind beide Arten in Verwendung:

##### Manuelle Lehrveranstaltungsaufzeichnung

Bei der manuellen Aufzeichnung erfolgt die Durchführung und Betreuung durch die Lehrenden, Tutoren/-innen oder spezielles Personal. Das hat den Vorteil, dass auf das Geschehen im Hörsaal Rücksicht genommen werden kann. So kann zum Beispiel bei einem Experiment oder bei Gebärdensprachdolmetschern/-innen eine Nahaufnahme erfolgen. Es kommen in den meisten Fällen handelsübliche und oft bereits vorhandene Komponenten wie zum Beispiel der Laptop des/der Vortragenden, eine digitale Videokamera oder Hörsaalmikrofone zum Einsatz. Dadurch

ist diese Methode in den meisten Fällen unabhängig von den räumlichen Gegebenheiten durchführbar. Somit können die Aufzeichnungen auch an beliebigen Orten erfolgen. So kann zum Beispiel eine Videoeinheit für einen Vorlesungsersatz im Büro des/der Professors/-in, in einem Arbeitszimmer oder einem Studio aufgezeichnet werden.

In den meisten Fällen wird die Aufnahme im Anschluss an die Aufzeichnung noch nachbearbeitet bevor, sie den Studierenden zur Verfügung gestellt wird. Durch den hohen Aufwand mit dem der Einsatz dieser Methode verbunden ist, ist eine Verwendung im großen Maße unpraktikabel und kostspielig. Sie wird daher nur in begründeten Einzelfällen oder unter besonderem Einsatz des/der Lehrenden eingesetzt (siehe [FHK08], [McC08], [Mer07, S: 19 - 33, 50 - 52]).

### **Automatische Lehrveranstaltungsaufzeichnung**

Zur einfachen und kostengünstigen Aufzeichnung möglichst vieler Lehrveranstaltungen werden heutzutage Systeme eingesetzt, die eine manuelle Betreuung und Nachbearbeitung minimieren. Friedland et. al. definiert dabei den Begriff folgendermaßen:

„The term automated lecture recording describes processes that aim to capture instructional classroom or lecture-hall events with a minimal use of manual processing“ [FHK08].

Die automatische Durchführung kommt auch der Anforderung der Lehrenden nach sehr einfachen und unaufdringlichen Lehrveranstaltungsaufzeichnungen entgegen. Es ist das Ziel dieser Systeme, Aufnahmen als „Nebenprodukt“ zum alltäglichen Lehrbetrieb zu erstellen:

„Automatic lecture recording basically aims to produce recorded lectures as a side-effect of a live session by being minimally invasive in the classroom event, not requiring constant supervision by humans, and automating the post-processing and archiving of the produced files as much as possible“ [FHK08].

Bei der automatischen Lehrveranstaltungsaufzeichnung kommen heute speziell dafür konzipierte Systeme zum Einsatz. Sie werden meist in den größeren und häufig genutzten Hörsälen installiert. Ergänzend dazu und um die Aufnahmen dennoch flexibel und unabhängig von den örtlichen Gegebenheiten durchführen zu können, werden oft auch mobile Einheiten verwendet. Infolge der Automatisierung ist es von Bedeutung, die wichtigsten Informationsquellen zu erfassen, die in der Regel in Vorlesungen eingesetzt werden. Es handelt sich dabei vor allem um den gesprochenen Inhalt. Der primäre Wissensvermittlungskanal wird fast immer durch Präsentationen, Multimediainhalte oder Experimente unterstützt. Sie sollten ebenfalls von den Lehrveranstaltungsaufzeichnungssystemen erfasst werden. Der inzwischen besonders häufige Einsatz von computergestützten Präsentationen kommt den Systemen sehr entgegen. Meist wird dazu direkt der Bildschirminhalt des Rechners aufgezeichnet. Viele Systeme bieten zusätzlich auch Schnittstellen zum Anschluss externer Kameras. Dadurch können Experimente, Tafelbilder, die vortragende Person oder aber auch Gebärdensprachdolmetscher/-innen erfasst werden. Somit ist es möglich, die Vorteile der manuellen mit der automatisierten Aufzeichnung zu kombinieren (siehe [FHK08], [Her11, S: 13 - 35], [McC08], [Mer07, S: 19 - 33]).

### 5.1.5 Wahrnehmung, Nutzungsverhalten und Auswirkung

Bei jeder neuen und optionalen Technik stellen sich rund um den Einsatz von Lehrveranstaltungsaufzeichnungen Fragen über die Nutzung, Auswirkungen und die dadurch bedingten Vor- und Nachteile. Im Zusammenhang damit steht das mögliche Fernbleiben der Studenten/-innen von Vorlesungen im Vordergrund. Zur Beantwortung wurden zahlreiche Experimente, Studien und Befragungen der Studierenden aber auch der Lehrenden durchgeführt. Bei deren Analyse ist besonders die Erfassung, wie die Aufzeichnungen durchgeführt aber auch in den Kursen eingesetzt wurden, wesentlich. Das hat mitunter einen sehr großen Einfluss auf die Ergebnisse. Bei den nachfolgend betrachteten Studien wurde die Lehrveranstaltungsaufzeichnung großteils als ergänzendes Material eingesetzt.

#### Studierende

Ein Großteil der durchgeführten Studien beschäftigt sich mit der Sicht der Studierenden auf die Aufzeichnungen und deren Auswirkungen auf die Studenten/-innen. Folgende Punkte lassen sich feststellen:

- **Allgemeine Bewertung**

Wie bereits eingangs erwähnt stehen die Studierenden den Lehrveranstaltungsaufzeichnungen meist sehr positiv gegenüber. In den meisten Studien werden die bereitgestellten Aufnahmen von einem Großteil der Studenten/-innen verwendet. Sie werden meist als wertvoll und für das Erlernen und Verstehen des Stoffes hilfreich bewertet (siehe [DCL09], [Her11, S: 40, 41], [Lar10], [SBG10], [Top11], [Wil06]).

Besonders positiv werden die Aufzeichnungen von fremdsprachigen und behinderten Studierenden aufgenommen (siehe [Cha07], [Lar10]). Im nachfolgenden Abschnitt 5.1.6 ab Seite 54 wird der Einsatz von Vorlesungsaufnahmen für behinderte Studierende ausführlich thematisiert.

- **Auswirkungen auf die Leistung der Studierenden**

Das ist der interessanteste und wichtigste Bereich bei der Verwendung von Lehrveranstaltungsaufzeichnungen. Die Studierenden geben an, dass sie Aufnahmen für das Erlernen des Stoffes als hilfreich bewerten. Entscheidend ist, wie die Aufzeichnungen eingesetzt werden. Werden sie als ergänzendes Material eingesetzt berichten Studien von einer gleichbleibenden Leistung (siehe [BA04], [Cha07]).

Im Gegensatz dazu hält Toppin fest, dass der Einsatz von Lehrveranstaltungsaufzeichnungen gute Voraussetzungen bietet, die Studentenleistung zu verbessern und somit die Kursabbrüche und negativen Abschlüsse zu vermindern (siehe [Top11]).

Werden die Aufzeichnungen dazu im Modell des „Inverted Classroom“ eingesetzt, konnten die Studierenden laut Talbert wesentlich bessere Leistungen erzielen und beherrschten den zu erlernenden Stoff besser (siehe [Tal12]).

- **Nutzung der Aufzeichnungen**

Neben der Auswirkung auf die Leistung der Studierenden ist es vor allem interessant, wie und für welchen Zweck die Lehrveranstaltungsaufzeichnungen eingesetzt werden. In

der von Hermann (siehe [Her11, S: 33 - 58]) durchgeführten Studie zeigte sich, dass die Studierenden die Aufnahmen für folgende Dinge verwenden (dabei waren Mehrfachnennungen erlaubt):

- 69,8% für die Bearbeitung von Übungsbeispielen.
- 66,6% als Ersatz für den Besuch der Vorlesung.
- 64,9% bei Verständnisproblemen.
- 49,7% zur Nachbearbeitung des Stoffes.
- 39,3% zum gezielten Nachschlagen von Informationen.

Die Ergebnisse decken sich größtenteils mit denen anderer Studien (siehe [BA04], [DCL09]). Besonders auffallend ist hierbei jedoch die Angabe, dass zwei Drittel der Befragten die Aufzeichnungen als Ersatz für den Besuch der Vorlesung nutzten. Diese Zahl lässt vorerst vermuten, dass durch den Einsatz der Aufzeichnungen ein Großteil der Studierenden den Vorlesungen fern bleibt. Im nächsten Abschnitt wird jedoch beschrieben, dass das meistens nicht der Fall ist. Aus der Befragung geht weiters hervor, dass es sich dabei sehr oft nur um einzelne Termine handelt. Das wird auch von anderen Studien bestätigt. Die Studierenden gaben an, die Aufzeichnungen als Ersatz für Vorlesungstermine zu verwenden, wenn sie aus familiären Gründen, Krankheit oder wegen ihrer Arbeit daran nicht teilnehmen konnten (siehe [Cha07], [Lar10]).

Bei Auswertungen, wann auf die Aufnahmen zugegriffen wird, zeigt sich, dass die Studierenden vermehrt vor Übungen, Abgabeterminen oder Prüfungen darauf zugreifen (siehe [Her11, S: 33 - 58]).

#### • Auswirkungen auf den Besuch von Präsenzterminen

Sehr viele Studien beschäftigen sich allein mit der Frage der Auswirkung von Lehrveranstaltungsaufzeichnungen auf den Besuch von Vorlesungen. Es zeigte sich bei fast allen untersuchten Publikationen, dass sich die Besuchszahlen nur geringfügig verringerten. Es gab auch Vorlesungen, bei denen sie sich merklich verringerten aber auch, dass sie in manchen Lehrveranstaltungen über den Verlauf des Semesters konstant hoch blieben. Bei den Befragungen wurde in den meisten Fällen die Einschätzung sowohl der Studierenden als auch der Lehrenden eingeholt. Manchmal wurden dabei auch Zählungen der Anwesenden durchgeführt (siehe [BA04], [DCL09], [Lar10], [SBG10]).

In der Regel wissen die Studenten/-innen, dass eine Aufzeichnung kein Ersatz für den Besuch der Vorlesung ist:

„Students in a focus group reported that they regarded the face-to-face sessions as being more valuable than the captures, saying that they were ‘not equal to lectures’; ‘I would not consider them to be a substitute for the live event. Because you can’t ask questions; there is no actual interaction’“ [DCL09].

Toppin hält dazu fest, dass das Thema sinkender Besuchszahlen wesentlich unproblematischer ist, wenn die Aufzeichnungen eine Voraussetzung für Aktivitäten während des Präsenztermins sind (siehe [Top11]).

- **Auswirkungen auf der Verhalten während der Vorlesungstermine**

Einige Autoren berichten von einer Verbesserung der Aktivität und Mitarbeit der Studierenden während der Präsenztermine. Die Begründung liegt meist darin, dass sie nicht gezwungen sind, Mitschriften anzufertigen, sondern dem Vortrag aktiv folgen können (siehe [BA04], [DCL09], [Top11]). Davis et al. fassen das folgendermaßen zusammen:

„Balfour (2006) stated that the availability of recorded lectures for review changes the way that students act in face-to-face sessions and ‘allows students to concentrate on the material being presented and to participate in any discussion that takes place.’ This is in contrast to students spending their time in class ensuring that all notes are fully comprehensive when the lecture is the one and only opportunity that they will have to hear the information. This argument is supported by Biggs (2003) who has stated that students struggle with comprehending the message of the lecture whilst simultaneously recording its gist through note-taking“ [DCL09].

## Lehrende

Im Gegensatz zu der Vielzahl an Studien über Studierende und Lehrveranstaltungsaufzeichnungen, gibt es relativ wenige, die sich mit den Lehrenden beschäftigen. Die Auswirkungen und Vorteile für sie werden meist allgemein festgehalten. Selten wurden die Vortragenden jedoch direkt befragt. Zum Thema Lehrveranstaltungsaufzeichnungen und deren Auswirkungen und Nutzungen aus Sicht der Lehrenden lassen sich dabei die folgenden Punkte zusammenfassen:

- **Allgemeine Bewertung**

In der Regel sehen die Lehrenden sehr schnell die möglichen Vorteile für die Studierenden. Hierbei ist besonders die Möglichkeit der verbesserten Zugänglichkeit zu den Lehrveranstaltungsinhalten für behinderte Studenten/-innen aber auch die Verwendung zur Nachbearbeitung hervorzuheben. Die Vortragenden erkennen jedoch selten einen Vorteil für sich selbst. Im besten Fall ist für sie die Aufzeichnung ohne Mehraufwand verbunden. Daher ist die Motivation diese einzusetzen oft sehr gering (siehe [Cha07]).

Wurde eine Vorlesung aufgenommen und die Vorbehalte der Lehrenden sind nicht eingetreten beziehungsweise haben sie auch Vorteile für sich selbst erkannt, wurde die Aufzeichnung im Allgemeinen als positiv empfunden (siehe [Lar10]).

- **Vorbehalte gegenüber Vorlesungsaufzeichnungen**

Rund um den Einsatz von Lehrveranstaltungsaufzeichnungen gibt seitens der Lehrenden einige Vorbehalte. Es wird meist das Ausbleiben der Studierenden in den Vorlesungen befürchtet (siehe [Cha07], [SBG10]). Es wurde bereits zuvor beschrieben, dass das meistens nicht der Fall ist. Diese Befürchtung kann verschiedene Hintergründe haben. Natürlich ist es nachvollziehbar, dass niemand gerne Vorträge in leeren Hörsälen hält. Ein weiterer Grund kann auch sein, dass die Vortragende Person einen Abfall der Studentenleistung durch das Fernbleiben befürchten. Dazu meint Larkin jedoch, dass der Besuch der Vorlesung nicht zwingend mit dem Lernerfolg verknüpft werden kann. Die Aufzeichnungen

sind für die Studierenden lediglich ein weiteres Mittel, um sich Wissen anzueignen und um nicht besuchte Vorlesungstermine nachzuholen beziehungsweise auszugleichen (siehe [Cha07], [Her11, S: 33 - 58], [Lar10]).

Weitere Vorbehalte gegenüber Lehrveranstaltungsaufzeichnungen sind, dass der eigene Vortragsstil nicht für die Aufzeichnung geeignet ist oder sie nicht in das Konzept des Kurses passt (siehe [Cha07]).

- **Nutzen der Aufzeichnung**

Der Einsatz von Aufnahmen hat jedoch nicht nur Vorteile für die Studierenden, er birgt auch Vorzüge für die Vortragenden. Wie in Rhetorikkursen oft Videoaufzeichnungen zur Verbesserung des Präsentationsstiles eingesetzt werden, können Lehrveranstaltungsaufnahmen zur Evaluierung und Verbesserung des eigenen Vortrages dienen. Auf lange Sicht können dadurch auch unklare Abschnitte selbst verbessert werden und somit die eigene Lehre verbessert werden (siehe [DCL09], [Lar10]).

Die Aufzeichnungen bieten auch die Möglichkeit, um weiteren Studierenden die zu vermittelnden Inhalte zugänglich zu machen. Dadurch eröffnet sich die Chance mehr Studenten/-innen zu erreichen aber auch mit ihnen in Kontakt zu treten (siehe [DCL09], [Lar10]). Zuletzt ist es durch den Einsatz von Lehrveranstaltungsaufzeichnungen auch möglich, die Zeit in der Vorlesung besser zu nutzen. Im einfachsten Fall trägt sie dazu bei, dass die Besucher/-innen der Vorlesungen, wie zuvor bereits beschrieben, aufmerksamer und konzentrierter sind. Auch der Einsatz von Konzepten wie dem auf Seite 46 zusammengefassten „Inverted Classroom“ können dazu beitragen.

- **Auswirkung auf das Verhalten während des Vortrages**

Der Einsatz von Aufzeichnungen in Lehrveranstaltungen hat auch Auswirkungen auf das Verhalten der Lehrenden während ihres Vortrages. Sie berichten davon, sich selbst zu zensurieren, so zum Beispiel private Informationen oder Erzählungen wegzulassen, weil sie wissen, dass sie gefilmt werden (siehe [Cha07]).

### 5.1.6 Lehrveranstaltungsaufzeichnungen für hörbehinderte Studierende

Bisher wurde die Lehrveranstaltungsaufzeichnung allgemein behandelt. Viele Studien und Publikationen betonen dabei die Vorteile, Chancen und Nutzen der Aufnahmen besonders für behinderte Studierende. Zugleich ergab die Literatursuche, die im Zuge der Diplomarbeit durchgeführt wurde, dass nur sehr wenige wissenschaftliche Veröffentlichungen das Thema der Vorlesungsaufzeichnungen speziell für behinderte beziehungsweise hörbehinderte Studenten/-innen behandeln. Die Technik wird jedoch definitiv zu deren Unterstützung eingesetzt (siehe [Echa], [McC08], [Roc09], [SAH10]).

Eine denkbare Erklärung dafür ist, dass davon nur ein Bruchteil aller Studierenden betroffen ist. Dennoch erkennen viele Lehrende die Vorteile sowie den möglichen Beitrag zur besseren Zugänglichkeit zu den Inhalten von Lehrveranstaltungen für behinderte Studenten/-innen (siehe [Cha07]).

Behinderte Personen sind je nach Art ihrer Einschränkung mit teils erheblichen Barrieren im Studium konfrontiert. Im österreichischen Bildungssystem erfolgt die Wissensvermittlung zu

einem Großteil durch die gesprochene Sprache. In Anbetracht des teilweisen oder vollständigen Hörverlustes ist das für hörbehinderte Studierende besonders problematisch:

„From these and other works, it has been established that the communications mechanism between the lecturer and the deaf student is one of the key areas that affect performance. In particular the lecture theatre environment is identified as an area where the information exchange is not particularly effective“ [SAH10].

Vorlesungen werden daher in der Regel von Simultandolmetschern/-innen in die Gebärdensprache übersetzt und somit den gehörlosen Studierenden eine Teilnahme an den Kursen ermöglicht. Die Wissensvermittlung und -Aufnahme erfolgt für sie durch den Sehsinn. Er ist ihr primärer Kommunikationskanal. Folglich ist es bei der Lehrveranstaltungsaufzeichnung wesentlich, die Gebärdensprachdolmetscher/-innen im Video möglichst gut erkennen und verstehen zu können. Das wird durch folgendes Zitat von McClure verdeutlicht. Es bezieht sich auf die Vorlesungsaufzeichnung für hörende Studierende, dennoch wird dadurch die Wichtigkeit der Aufnahme des primären Kommunikationskanals der jeweiligen Zielgruppe verdeutlicht:

„While video isn't a must-have, "audio is extremely important," [...] "If the audio is not adequate, it reduces the usefulness." [...]“ [McC08].

Neben den zuvor erwähnten ÖGS Dolmetschern/-innen ist der Einsatz von Untertiteln eine weitere Möglichkeit, um den gesprochenen Vortrag für hörbehinderte Studierende zugänglich zu machen. Sie können wie in Kapitel 2.2 auf Seite 9 bereits beschrieben entweder in Echtzeit direkt in der Vorlesung eingesetzt werden oder nachträglich einer Lehrveranstaltungsaufzeichnung hinzugefügt werden.

Sind die Inhalte der Aufnahmen zugänglich, bieten die Aufzeichnungen zumindest die zuvor beschriebenen Vorteile für sie. Darüber hinaus stellen sie für behinderte Studierende einen wesentlichen Beitrag zur Zugänglichkeit der Inhalte von Kursen dar. Die Publikation von Larkin behandelt die Auswirkungen, Vor- und Nachteile der Lehrveranstaltungsaufzeichnung allgemein, dennoch lässt sich das nachfolgende Zitat besonders auf behinderte Studenten/-innen anwenden. Schließlich ist es ihnen oftmals aufgrund ihrer Behinderung nicht möglich, die Vorlesungen zu besuchen oder deren Inhalte problemlos aufzunehmen.

„The provision of recorded lectures can enhance teaching and learning opportunities beyond mere convenience for those students who are otherwise unable to attend lectures“ [Lar10].

Für hörbehinderte Studierende stellen die folgenden Punkte einen erheblichen Schritt zum barrierefreien Zugang zu Lehrveranstaltungsinhalten dar:

- **Möglichkeit zur Nachbearbeitung des Stoffes**

Bei Gebärdensprachen handelt es sich um visuelle Sprachen. Das bedeutet die Kommunikation erfolgt über den Sehsinn und erfordert somit auch Blickkontakt. Das ist besonders in Lehrveranstaltungen problematisch. Richtet der/die gehörlose Student/-in seine Augen auf etwas anderes als die Dolmetscher/-innen also zum Beispiel auf die Folien, den Vortragenden oder die eigenen Notizen, so versäumt er/sie zwangsläufig einen Teil des Inhaltes

der Vorlesung. Für hörende Studierende besteht dieses Problem nicht, denn sie können dem Vortrag des Lehrenden auch weiterhin folgen, wenn sie gerade die Folien betrachten. Hierbei stellen die Aufzeichnungen von Lehrveranstaltungen eine große Abhilfe dar. Dadurch können die hörbehinderten Studenten/-innen den Stoff wesentlich besser aufarbeiten und so zum Beispiel das Video pausieren oder nochmals abspielen, um Folien oder ein Experiment zu betrachten.

- **Entlastung beim Niederschreiben von Notizen**

Die im vorhergehenden Punkt beschriebene Problematik bringt jedoch auch noch einen zweiten großen Nachteil für gehörlose Studierende mit sich. Es ist für sie besonders schwierig, während des Vorlesungsbesuches Notizen anzufertigen. Dadurch sind die Videoaufzeichnungen - besonders der Gebärdensprachdolmetscher/-innen - für gehörlose Studenten/-innen ein großer Beitrag, um bestehende Barrieren abzubauen. Ein Professor der Gallaudet University, einer Universität in den USA speziell für gehörlose Personen, in der in American Sign Language (ASL) unterrichtet wird, beschreibt dies folgendermaßen:

„Those who use it have told me it has been a real life saver. As deaf students, they need to be looking at the professor throughout the entire lecture; they can't be taking notes. [...] It helps them learn more effectively because they can concentrate on me during the lecture. They don't have to split their time anymore during lectures; they just focus on what is going on in class“ [Echa].

- **Möglichkeit zur Transkription des Inhaltes**

Ist eine Vorlesung aufgezeichnet, kann der gesprochene Inhalt leichter durch eine Spracherkennungssoftware aber auch mit manuellen Erzeugungsmethoden transkribiert werden. Der Text kann den hörbehinderten Studierenden entweder direkt zur Verfügung gestellt werden oder unter Zuhilfenahme spezieller E-Learning-Plattformen, wie zum Beispiel Synote (siehe kommender Abschnitt), mit der Aufzeichnung synchronisiert werden.

### **Bewertung von Lehrveranstaltungsaufzeichnungen durch behinderte Studierende**

Im Zuge der Literaturrecherche wurde lediglich eine Studie über den Einsatz und den Nutzen von Lehrveranstaltungsaufzeichnungen für behinderte Studierende gefunden. Sie wurde an der University of Western Australia durchgeführt. Dabei beantworteten 21% der 642 Studenten/-innen den Fragebogen, bei dem sich ein klares Bild zeigte (siehe [Wil06]):

- 16% der befragten Studierenden verwendeten die zur Verfügung gestellten Lehrveranstaltungsaufzeichnungen immer, 40% regelmäßig und 33% gelegentlich.
- Nur ein sehr geringer Prozentsatz verwendete spezielle Software, um auf die Aufzeichnungen barrierefrei zuzugreifen (3% verwendeten eine ASR, 1% einen Screenreader und 1% eine Bildschirmlupe).
- 43% der Befragten besuchten alle Vorlesungstermine und 45% besuchten sie regelmäßig.
- 65% der Studierenden bewerten die an der Universität eingesetzte Lehrveranstaltungsaufzeichnungen als wesentlich. 32% gaben an, dass sie zumindest hilfreich ist.

- 73% verwendeten die Aufzeichnungen, um nicht verstandene Inhalte zu wiederholen.
- 36% aller befragten Studierenden war dabei nicht in der Lage, während der Vorlesung Notizen aufzuschreiben.

Bei der Befragung gaben viele behinderte Studenten/-innen an, dass die Aufzeichnungen einen wichtigen Beitrag zum Erlernen der Inhalte darstellen:

„Many noted that the system has been vital in enabling their learning“ [Wil06].

## 5.2 Synote

Im ersten Teil dieses Kapitels wurde bereits des öfteren die Wichtigkeit der E-Learning-Umgebung betont, in der die Lehrveranstaltungsaufzeichnungen eingebunden werden. Im vorliegenden Abschnitt wird eine Plattform zu besserer Nutzung bestehender Aufnahmen vorgestellt. Synote soll es ermöglichen, diese mit anderen Informationsquellen zusammenzuführen, einfach zu benutzen und zu durchsuchen. Somit sollen die Inhalte für Lernende besser zugänglich werden (siehe [WWM<sup>+</sup>09], [Wal10]). Das Prinzip hinter Synote ist in Abbildung 5.1 dargestellt. Bei der Entwicklung an der University of Southampton wurde speziell auf hörbehinderte Studierende Rücksicht genommen. Weiters ist die Plattform „unter Einhaltung von Urheberrechtsbestimmungen für Bildungseinrichtungen ohne Lizenzkosten“ [Hat11, S: 134] einsetzbar.

Synote ermöglicht es, externe Audio- und Videoaufzeichnungen mit deren Transkript und den Präsentationsfolien zu verknüpfen und synchron abzuspielen. Bei der Wiedergabe wird die aktuelle Stelle im Text und den Folien farblich gekennzeichnet (siehe Abbildung 5.2). Darüber hinaus ist es den Benutzern/-innen möglich, Anmerkungen, Notizen und Kommentare zu erstellen. Sie werden ebenfalls einer bestimmten Stelle in der Aufzeichnung zugewiesen (siehe [WWM<sup>+</sup>09], [Wal10]).

Durch die Einbeziehung der E-Learning-Umgebung ergeben sich völlig neue Möglichkeiten bei der Nutzung von Aufzeichnungsdokumenten. Durch deren Einsatz ist es möglich, technische Limitierungen bestehender Wiedergabesysteme für Vorlesungsaufzeichnungen zu umgehen, so zum Beispiel, wenn es damit nicht möglich ist, die Videos mit Untertiteln zu versehen. Das stellt für hörbehinderte Studierende allgemein und besonders für die Teilnehmer/-innen des GESTU Projekts eine große Chance dar.

Die Software wird ständig weiterentwickelt und um neue Funktionalitäten erweitert. Eine der letzten Neuerungen ist das verteilte Korrigieren von Fehlern in Transkripten (siehe [WL12]). Eine detaillierte Diskussion von Synote und dessen Einsatz für hörbehinderte Studierende erfolgt von Hattinger den Diplomarbeiten (siehe [Hat11, S: 133 - 137], [Hat13]).

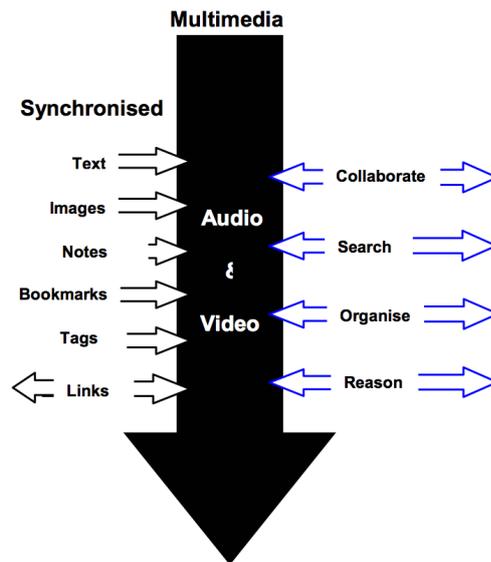


Abbildung 5.1: Das Prinzip von Synote [WWM<sup>+</sup>09, Figure 1].

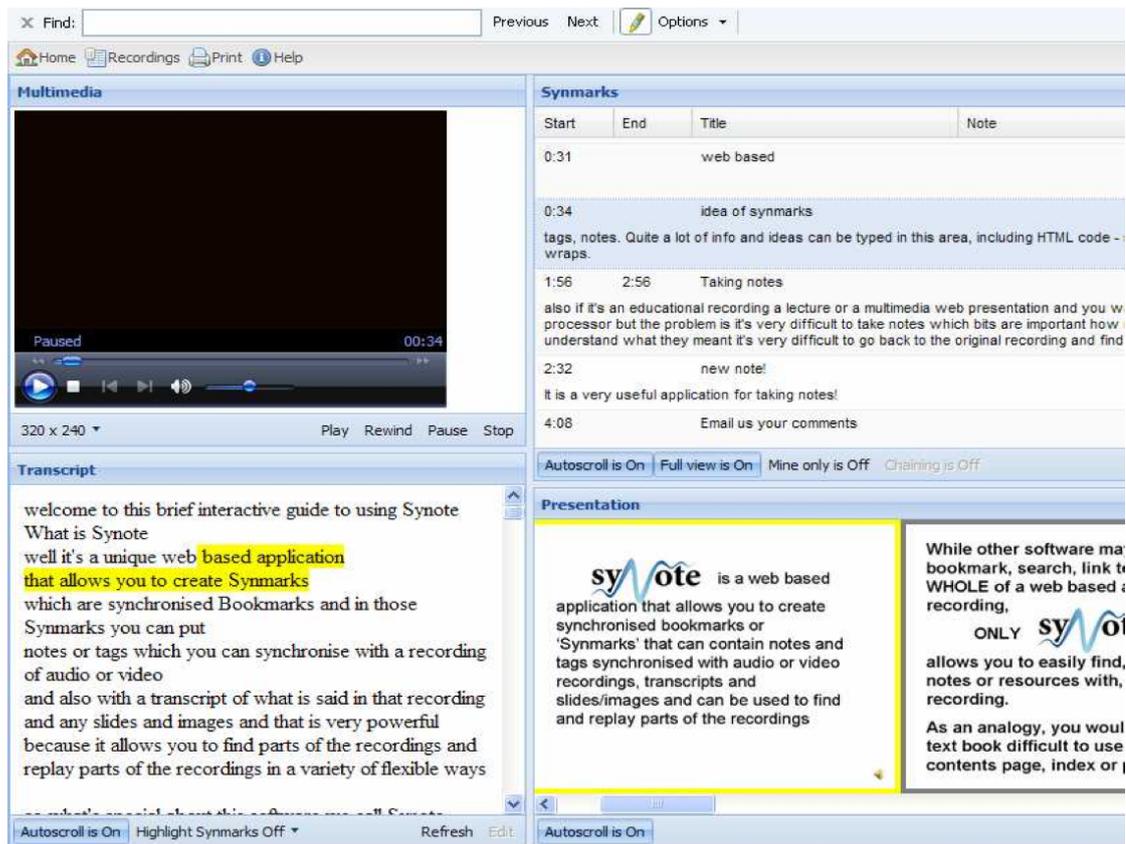


Abbildung 5.2: Das Synote Benutzerinterface [WWM<sup>+</sup>09, Figure 3].

## 5.3 Lehrveranstaltungsaufzeichnungen im GESTU Projekt

Die Lehrveranstaltungsaufzeichnung für die hörbehinderten Studierenden des GESTU Projekts war eines der ersten aber auch am meisten eingesetzte technische Hilfsmittel. Bereits im ersten Semester des Modellversuchs, im Herbst 2010, wurde mit der Aufnahme universitärer Vorlesungen begonnen. Das geschah auf Basis der von Hattinger im Zuge seiner ersten Diplomarbeit ausgesprochenen Empfehlung zum Einsatz der E-Learning-Plattform Synote (siehe [Hat11, S: 133 - 137]). Dadurch sollte ein Beitrag zur Erleichterung des Studienalltags der hörbehinderten Projektteilnehmer/-innen durch die zuvor beschriebenen Möglichkeiten und Vorteile von Lehrveranstaltungsaufzeichnungen geschaffen werden. Es wurde vor allem die Einbindung der Videoaufnahmen und deren Transkripte in Synote angestrebt.

Durch die in Abschnitt 3.2 auf Seite 24 beschriebene Kooperation mit dem EML zum Training des von ihnen entwickelten sprecherunabhängigen Spracherkennungssystems für freie Sprache, eröffnete sich die Aussicht auf eine mögliche automatische Erstellung der Transkripte. Hierfür diente vor allem der von Wald beschriebene Einsatz von Synote und der automatischen Erstellung von Untertiteln für hochgeladene Audio- und Videomitschnitte von Vorlesungen als Vorbild (siehe [Wal10]). Um durch die EML ASR gute Ergebnisse für den, an den Wiener Universitäten gesprochenen österreichischen Dialekt zu erzielen, musste sie mit möglichst vielen Audiomitschnitten aus Vorlesungen trainiert werden. Es war jedoch nicht nur die Menge der Trainingsdaten, sondern auch die Berücksichtigung einer großen Vielfalt an wissenschaftlichen Fachrichtungen für das Training entscheidend. Deren Sammlung stellte daher eine zusätzliche Motivation für die Durchführung der Vorlesungsaufzeichnungen dar.

Bedingt durch die fehlende Literatur speziell zum Thema der Lehrveranstaltungsaufzeichnung für hörbehinderte Studierende, handelte es sich bei den Aufzeichnungen im Zuge des Modellversuchs um einen Pilotversuch. Ziel war, den Beitrag dieser Technologie zur Erleichterung des Studienalltags für diese Zielgruppe zu ergründen und evaluieren. Durch die positiven Rückmeldungen der Studenten/-innen nach dem ersten Semester wurde der Einsatz der technischen Maßnahme noch weiter ausgedehnt. Es konnte daher die Anzahl der zukünftig aufgezeichneten Vorlesungen erhöht werden. Weiters wurden auch die von Professor Dotter veranstalteten Gebärdensprachlinguistik-Lehrveranstaltungen auf der TU Wien zum Teil unter Verwendung von zwei Kameras aufgezeichnet.

### 5.3.1 Durchführung

Bei der Planung und der Durchführung der Lehrveranstaltungsaufzeichnungen wurde besonders auf die Bedürfnisse, der am Modellversuch teilnehmenden, hörbehinderten Studierenden geachtet. Daher wurden ausschließlich Video- und keine reinen Audioaufzeichnungen eingesetzt. Es wurde dabei auf die gute Sichtbarkeit der ÖGS Dolmetscher/-innen und die Verständlichkeit der Übersetzung in die Gebärdensprache geachtet (siehe Abschnitt „Die Aufzeichnung in den Vorlesungen“ auf Seite 61). Zusätzlich wurde bei allen Aufnahmen auch der gesprochene Vortrag in möglichst guter Tonqualität aufgezeichnet, denn sie wurden allen Kursbesuchern zur Verfügung gestellt, sofern die vortragende Person dem zustimmte. Außerdem wurden die Audiomitschnitte der gesprochenen Vorträge auch zum Training der EML ASR verwendet (siehe Abschnitt 3.2 auf Seite 24).

Aufgrund der zuvor beschriebenen Aufnahme der Gebärdensprachdolmetscher/-innen wurden alle Lehrveranstaltungsaufzeichnungen manuell durchgeführt. Da es sich beim Modellversuch um ein Universitäten übergreifendes Projekt handelt, wurden verschiedene Vorlesungen in den unterschiedlichen Hörsälen an den Wiener Universitäten aufgezeichnet. Es wurden keine eigenen und speziellen Aufzeichnungssysteme eingesetzt, sondern es wurde ausschließlich auf mobile Lösungen unter Zuhilfenahme bereits vorhandener Ressourcen der jeweiligen Bildungseinrichtungen gesetzt. Es wurden daher auch keine speziellen Trackingsysteme zum Einrichten von Kameras auf die Dolmetscher/-innen verwendet. Die Aufnahmen selbst wurden von Tutor/-innen des Projekts betreut.

Während des GESTU Projekts wurde daher mit den für die Lehrveranstaltungsaufzeichnungen an den jeweiligen Universitäten zuständigen Abteilungen zusammengearbeitet. Durch diese Kooperation konnten bestehende Infrastrukturen und Ressourcen verwendet werden. Somit mussten keine speziellen Systeme zur Lehrveranstaltungsaufzeichnung aber auch keine Streaming- oder Downloadmöglichkeiten durch das GESTU Projekt etabliert werden. Die Aufnahmen des Modellversuchs konnten dadurch mit geringen Kosten durchgeführt werden.

### **Rechtliche Aspekte zur Aufzeichnung von Lehrveranstaltungen**

Laut §14 des österreichischen Urheberrechtsgesetzes liegt das Verwertungsrecht eines Werkes beim/bei der Urheber/-in. Im Falle einer Vorlesung handelt es sich dabei laut §2 des Urheberrechtsgesetzes um ein sogenanntes Sprachwerk. Die vortragende Person muss daher der Aufzeichnung zustimmen. Da es sich bei Videoaufnahmen um personenbezogene Daten handelt, muss aus Gründen des Datenschutzes die Zustimmung aller auf dem Video gezeigten Personen erfolgen (siehe [Hat11, S: 33 - 35]).

Vor der Durchführung von Lehrveranstaltungsaufzeichnungen des GESTU Projekts musste daher die Zustimmung der vortragenden Person aber auch der Dolmetscher/-innen eingeholt werden.

### **Ablauf zu Semesterbeginn**

Wie bereits im Abschnitt „Ablauf zu Semesterbeginn“ des Kapitels „Respeaking im GESTU Projekt“ auf Seite 34 beschrieben, wurden sämtliche technische Maßnahmen vor dem Semesterbeginn koordiniert. Dabei war der Ablauf für die Einteilung des Respeakings aber auch der Lehrveranstaltungsaufzeichnung ident. Die am Modellversuch teilnehmenden Studierenden gaben die Lehrveranstaltungen bekannt, die sie im kommenden Semester besuchen werden und welche für sie aufgezeichnet werden sollen.

Die vortragenden Personen mussten, wie zuvor beschrieben, den Aufnahmen ihrer Vorlesungen zustimmen. Daher wurden sie von der Behindertenbeauftragten der jeweiligen Universität stellvertretend für das GESTU Projekt um ihr Einverständnis für die Durchführung der Aufzeichnung ersucht. Weiters wurde auch erörtert, ob sie nur den hörbehinderten Studierenden oder allen am Kurs teilnehmenden Studenten/-innen zur Verfügung gestellt werden. Im ersten Fall wurden die Videos durch ein Passwort geschützt. Außerdem wurde dabei um die Zustimmung zur Verwendung der Audioaufzeichnung des gesprochenen Vortrages zum Training der EML ASR ersucht (siehe Abschnitt 3.2 auf Seite 24).

Sofern der/die Lehrende mit der Aufnahme seiner/ihrer Vorlesung einverstanden war, übernahmen Tutoren/-innen des GESTU Projekts die Durchführung und deren Betreuung in den Lehrveranstaltungen. Die Suche und Einteilung der sogenannten eTutoren wurde ebenfalls noch vor dem Semesterbeginn durchgeführt. Zur Aufwandsabschätzung diente dieselbe Formel, wie sie auch für die Einteilung des Respeakings verwendet wurde. Sie ist nachfolgend abgebildet. Dabei steht die „Anzahl“ für die Anzahl sämtlicher Lehrveranstaltungstermine und die „Dauer“ für die Dauer eines Termins. Die Schätzung umfasst die gesamten Tätigkeiten, von der Aufzeichnung bis zu Nachbearbeitung durch den/die eTutor/-in. Die Schätzung deckte sich sehr gut mit dem tatsächlichem Stundenaufwand während des Semesters.

$$\text{Aufwand zur Betreuung einer Lehrveranstaltung} = \text{Anzahl} * \text{Dauer} * 2 \quad (5.1)$$

Die Tutoren/-innen selbst wurden vor und während der ersten Lehrveranstaltungstermine in das verwendete Equipment, die Software und die Durchführung in den Vorlesungen eingeschult.

### Die Aufzeichnung in den Vorlesungen

Sämtliche Lehrveranstaltungsaufzeichnungen wurden manuell durchgeführt. Sie wurden von Tutoren/-innen des GESTU Projekts betreut. Somit entstand für die Vortragenden, wenn überhaupt, nur ein geringer Mehraufwand durch die Aufnahmen. Darüber hinaus fungierten die eTutoren/-innen als Ansprechperson für die Lehrenden, die Dolmetscher/-innen und die hörbehinderten Studierenden.

Die Tutoren/-innen waren dabei für die Mitnahme und das Funktionieren des verwendeten Equipments verantwortlich. Dabei handelte es sich um eine digitale Videokamera inklusive eines Stativs, ein Funkmikrofonset inklusive Reserveakkus sowie diverse Kabel. Eine genaue Beschreibung sowie die Anforderungen an diese Geräte ist am Ende des Abschnitts auf Seite 64 zu finden. Die Kamera wurde vor dem Start der Vorlesung vom Tutor/-in aufgebaut und falls benötigt die Sendeeinheit des Funkmikrofons der Vortragenden Person übergeben.

Um eine möglichst gute Verständlichkeit der Aufnahme der Übersetzung in die Gebärdensprache zu erreichen sind die nachfolgenden Punkte zu berücksichtigen:

- Möglichst große Aufzeichnung der Dolmetschenden Person.
- Die Kamera soll dabei auf Augenhöhe des/der Dolmetschers/-in aufgebaut werden.
- Frontale Aufnahme.
- Möglichst einfarbiger und zur Hautfarbe kontrastreicher Hintergrund.

Gemäß den oben angeführten Punkten ist eine gute Platzwahl für das Aufstellen der Videokamera wesentlich. Weiters ist sie auch von der Entscheidung, welche Informationsquellen der Vorlesung ebenfalls aufgezeichnet werden, abhängig. Es handelte sich dabei um den/die Dolmetscher/-in, die Vortragende Person, das Tafelbild und das Beamerbild. Durch die Aufnahme mehrerer Quellen wird ein umfassenderes Bild des Vortages vermittelt. Sofern in der Vorlesung nicht zwei Videos aufgezeichnet werden und im Anschluss daran synchron wiedergegeben



(a) Nahaufnahme einer Dolmetscherin



(b) Aufzeichnung einer Dolmetscherin gemeinsam mit dem Vortragenden

**Abbildung 5.3:** Zwei der gebräuchlichsten Bildeinstellungen für die Videoaufzeichnung von Lehrveranstaltungen für hörbehinderte Studierende, die im Zuge des GESTU Projekts durchgeführt wurden.

werden können, ist die zusätzliche Aufzeichnung der anderen Informationsquellen jedoch wegen der möglichst großen Aufnahme der Dolmetscher/-innen nur bedingt möglich. Daher wird mit dem/der anwesenden hörbehinderten Studenten/-in abgesprochen, was alles aufgezeichnet werden soll. Die meisten bevorzugten eine alleinige und möglichst große Aufzeichnung des/der Gebärdensprachdolmetschers/-in. Zwei der gängigsten Bildeinstellungen sind in Abbildung 5.3 zu sehen. Eine Gewichtung der einzelnen Informationsquellen der Vorlesung, die durch die Befragungen der hörbehinderten Studierenden erhoben wurde, ist im folgenden Abschnitt 5.3.2 auf Seite 66 zu finden. Vor dem Beginn der Vorlesung erfolgte eine Absprache mit den anwesenden Dolmetschern/-innen sowie dem/der Lehrenden, um möglichst ideale Bedingungen für die Aufzeichnung zu schaffen. In den meisten Fällen handelte es sich dabei um die Positionierung der einzelnen Personen. Während des Vortrages beaufsichtigte der/die anwesende Tutor/-in die Aufzeichnung. Dazu zählte zum Beispiel die Korrektur des Kamerabildes. Das ist notwendig, falls die Gebärdensprachdolmetscher/-innen, zumeist beim Wechsel, ihre Position ändern. Manchmal ist es auch notwendig, den Kamerafokus auf ein im Hörsaal gezeigtes Experiment zu wechseln. Weiters stellte er/sie die Funktionstüchtigkeit des Equipments sicher. In diesem Zusammenhang war besonders die Stromversorgung der einzelnen Geräte wesentlich, da diese in den meisten Fällen durch Akkumulatoren erfolgte.

Zur Durchführung und Betreuung der Lehrveranstaltungsaufzeichnungen in den Lehrveranstaltungen stand den Tutoren/-innen ein Leitfaden als Hilfestellung zur Verfügung. Dieser ist im Anhang (Kapitel 9.2.2 auf Seite 115) zu finden.

### **Lehrveranstaltungsaufzeichnung mit zwei Videokameras**

Neben der zuvor beschriebenen Lehrveranstaltungsaufzeichnung mit einer Kamera wurden während des GESTU Projekts auch Versuche mit dem Einsatz von zwei Videokameras durchgeführt. Diese konnten mit der Streaming-Plattform der TU Wien synchron abgespielt werden. Das wurde in den beiden Vorlesungen von Professor Dotter über die Gebärdensprachlinguistik und Gebärdensprachentwicklung eingesetzt. Er gestaltete seinen Vortrag sehr interaktiv. Aus diesem Grund fanden durch die Studierenden sehr viele Diskussionen in der ÖGS statt. Zu Beginn der Aufnahmen wurde nur eine Kamera eingesetzt. Bei Wortmeldungen mussten die Studenten/-innen vor sie treten. Durch die Rückmeldungen stellte sich schnell heraus, dass dadurch der Diskussionsfluss stark gehemmt wurde. Daher wurden Versuche mit zwei Videokameras durchgeführt. Das erlaubte die gleichzeitige Aufnahme der Übersetzung des gesprochenen Vortrages der Dolmetscher/-innen und der Beiträge der Studierenden.

Die beiden Videos wurden im Anschluss vom/von der Tutor/-in so geschnitten, dass deren Beginn übereinstimmt. Das war in den meisten Fällen problemlos möglich, benötigte aber natürlich insgesamt mehr Zeit. Abschließend wurden die Aufzeichnungen in die Streaming-Plattform der TU Wien eingebunden und konnten so synchron abgespielt werden.

### **Nachbearbeitung und Onlinestellen der Aufzeichnungen**

Im Anschluss an die Aufzeichnung wird die notwendige Nachbearbeitung der Videos und das Hochladen auf die Streaming-Plattform der jeweiligen Universität, ebenfalls vom/von der eTutor/-in, durchgeführt. Er/sie fügt dabei die einzelnen Videodateien zusammen, schneidet eventu-

elle überflüssige oder von der vortragenden Person gewünschte Stellen heraus und konvertiert es schlussendlich in ein für den Streamingserver geeignetes Format. Für sämtliche Abläufe stand wieder ein Leitfaden als Hilfestellung zur Verfügung. Die Anleitung zur Nachbearbeitung der Videos unter Zuhilfenahme der nachfolgend angeführten Software ist im Anhang (Kapitel 9.2.3 auf Seite 121) zu finden.

### **Vorhandene Infrastruktur an den Universitäten**

Alle im Zuge des GESTU Projekts aufgezeichneten Lehrveranstaltungen fanden an der Technischen Universität Wien oder der Universität Wien statt. Durch Kooperationen mit beiden Bildungseinrichtungen war es möglich, die vorhandenen Infrastrukturen zu verwenden. Daher musste die Bereitstellung der Videos nicht durch den Modellversuch gelöst werden. Da beide Einrichtungen unterschiedliche Lehrveranstaltungsaufzeichnungs- und Streaming-Systeme einsetzen, ergaben sich jedoch auch für die Durchführung und Gestaltung der Aufzeichnungen unterschiedliche Möglichkeiten:

- **Universität Wien**

Mit dem in den größten und am häufigsten verwendeten Hörsälen der Universität Wien installierten automatischen Lehrveranstaltungsaufzeichnungssystem ist es möglich, das Bild des Präsentationsrechners und den gesprochenen Vortrag aufzunehmen. Zusätzlich ist bei manchen dieser Anlagen auch eine Kamera zur Aufzeichnung des Vortragenden installiert. Bei Versuchen mit dem System, die Dolmetscher/-innen zu filmen, war es jedoch nicht möglich, den Bildausschnitt gut einzustellen.

Die Aufzeichnungsdokumente wurden den Studierenden durch einen selbsterstellten Streaming-Server zur Verfügung gestellt. Er unterstützt auch das Einspielen externer Filme. Während der Zeit des ersten GESTU Projekts war es jedoch nur möglich, ein Video gleichzeitig abzuspielen.

- **Technische Universität Wien**

An der TU Wien sind ebenfalls einige Hörsäle mit einem automatischen Lehrveranstaltungsaufzeichnungssystem ausgestattet. Es zeichnet gleichfalls das Bild des Präsentationsrechners in Kombination mit dem gesprochenen Vortrag auf. In die in den Hörsälen installierten Einheiten können weiters externe Videosignale eingespielt und somit mitaufgezeichnet werden. Dadurch konnte eine Videokamera angeschlossen werden, mit der die Dolmetscher/-innen gefilmt wurden.

Die einzelnen Aufzeichnungsdokumente wurden durch den Opencast Matterhorn<sup>1</sup> Streaming-Server zur Verfügung gestellt. Er ermöglicht auch das parallele Abspielen zweier Videos. Es ist auch möglich, manuell erstellte Videos in das System einzuspielen.

### **Eingesetztes Equipment**

Die frühe Kooperation mit der Universität Wien zu Beginn des Modellversuchs ermöglichte die Nutzung nicht nur der bestehenden Streaming-Infrastruktur, sondern auch des vorhandenen

<sup>1</sup> <http://opencast.org/matterhorn> - letzter Zugriff am 06.03.2013.

Aufzeichnungsequipments. Es handelte es sich um eine digitale Videokamera mit einem Eingang für ein externes Mikrofon, ein „Sennheiser ew100 G3“ Funkmikrofon Set und ein kleines Stativ zum Aufstellen der Kamera. Da das Set ab dem Wintersemester 2011 wieder in vermehrtem Ausmaß durch die Universität Wien verwendet wurde, musste ein eigenes Aufnahmeset für das GESTU Projekt angeschafft werden. Das geschah nicht nur zur weiteren Durchführung der Lehrveranstaltungsaufzeichnungen, sondern es wurde auch für die Fachgebärdenworkshops und Gebärdenaufnahmen verwendet. Das angeschaffte Funkmikrofon wurde darüber hinaus auch für die Live-Untertitelung mittels der Respeaking Technik eingesetzt. Es wurden folgende Komponenten gekauft, welche die nachstehenden Anforderungen erfüllen mussten:

- **Videokamera - Panasonic HDC-SD99EG-K**

Die Bildqualität aktueller digitaler Videokameras reicht bei weitem für die Lehrveranstaltungsaufzeichnung aus. Die Videos werden in den meisten Fällen im Internet als Stream oder zum Download angeboten und es kommen hierbei meistens keine sehr großen Bildauflösungen zum Einsatz.

Bedingt durch die räumliche Situation in Hörsälen und um möglichst optimale Aufnahmen der Gebärdensprachdolmetscher/-innen anfertigen zu können, ist es wesentlich, dass das Objektiv der Kamera gute Zoom- und Weitwinkelleigenschaften besitzt.

Bei den Vorlesungsaufzeichnungen im Zuge des GESTU Projekts musste auch der gesprochene Vortrag in möglichst guter Qualität aufgezeichnet werden. Daher musste die Kamera einen Eingang für ein externes Mikrofon besitzen. Beim gewählten Modell handelt es sich um einen 3,5mm Klinke Eingang.

Ein weiteres Entscheidungskriterium war das Speichermedium der Kamera. Es sollte möglichst viel Speicherplatz bieten, kostengünstig, wiederverwendbar und austauschbar sein. Die Wahl fiel daher auf SD (Secure Digital) Memory Cards.

Weiters war eine Kamera auszuwählen, welche die zuvor genannten Kriterien erfüllt aber nicht zu teuer ist. Die Wahl fiel auf die „Panasonic HDC-SD99EG-K“.

- **Funkmikrofon - Sennheiser ew100 G3, E-Band**

Manche Lehrveranstaltungsaufzeichnungen wurden allen Kursteilnehmern/-innen zur Verfügung gestellt oder für das Training der EML ASR verwendet (siehe Abschnitt 3.2 auf Seite 24). Es wurde daher auch der gesprochene Vortrag mit einem externen Mikrofon in guter Qualität aufgezeichnet. Dazu wurde ebenfalls das bereits im Abschnitt „Eingesetztes Equipment“ des Kapitels „Respeaking im GESTU Projekt“ auf Seite 36 beschriebene „Sennheiser ew100 G3“ Funkmikrofonset verwendet.

- **Kamerastativ - Velbon CX-Mini Stativ**

Es wurde ein kleines und kompaktes Stativ ausgewählt, das auf kleinen Hörsaaltischen aufgestellt werden konnte. Der Stativkopf des Modells kann in sämtliche Richtungen bewegt werden.

- **Diverse Kabel**

Für lang dauernde Aufnahmen lag dem Aufnahmeset ein 5m Strom-Verlängerungskabel bei. Darüber hinaus wurde für den Anschluss der Videokamera an das Lehrveranstaltungsaufnahmesystem der TU Wien ein 10m Clinch Kabel angeschafft.

### Eingesetzte Software

Infolge der manuellen Durchführung der Aufnahmen und den nicht in allen Hörsälen der Universitäten vorhanden speziellen Systemen zur Lehrveranstaltungsaufzeichnung war es oftmals nötig, die Videos nachzubearbeiten, bevor sie den Studierenden zur Verfügung gestellt werden konnten. Die Nachbearbeitung umfasste das Zusammenfügen mehrerer Dateien und die Konvertierung. Ersteres ist erforderlich, da viele digitale Videokameras die Aufnahmen nach einer bestimmten Zeit oder nachdem die Datei eine gewisse Größe erreicht hat, teilen. Die Nachbearbeitung wurde ebenfalls von den eTutoren/-innen durchgeführt. Bei der Auswahl der Software war es daher wesentlich, Programme zu wählen, welche die geforderte Funktionalität erfüllen, einfach in der Handhabung sind und gleichzeitig auch möglichst kostengünstig sind. Um die Videos im Apple OS X Betriebssystem nachzubearbeiten, wurde die folgende Software verwendet:

- **Turbo.264 HD<sup>2</sup>**

Das oben angeführte Programm wurde zur Nachbearbeitung der Videos unter dem Betriebssystem Apple Mac OS X verwendet. Damit konnten alle drei geforderten Nachbearbeitungsschritte unter Verwendung von nur einem Programm ausgeführt werden.

Für Microsoft Windows konnte zum Zeitpunkt der Suche keine Software gefunden werden, die kostengünstig ist beziehungsweise alle geforderten Schritte beherrscht. Es wurden daher die folgenden beiden gratis verfügbaren Programme zur Nachbearbeitung gewählt:

- **TSSnipper<sup>3</sup>**

Mit der „TSSnipper“ Software können die Ausgabedateien der von GESTU eingesetzten Kamera geschnitten werden, ohne sie davor oder während des Schneidens konvertieren zu müssen.

- **Any Video Converter<sup>4</sup>**

Die einzelnen Videodateien wurden im Anschluss daran mit dem „Any Video Converter“ zusammengefügt, komprimiert und in die benötigten Formate konvertiert. War kein Schnitt notwendig, wurde von den Tutoren ausschließlich das Programm verwendet.

### 5.3.2 Erkenntnisse

Bei der Lehrveranstaltungsaufzeichnung handelt es sich um das im GESTU Projekt am meisten eingesetzte technische Hilfsmittel. Es gab, wie auch beim Respeaking, bei den ersten Einsätzen eine Etablierungsphase. Ab der zweiten Hälfte des Modellversuchs war die Organisation und die Verwendung reibungslos möglich. Während der Eingangsphase wurden Arbeitsprozesse erarbeitet, um die Aufzeichnungen reibungslos zu organisieren und durchzuführen. Hierbei

<sup>2</sup> <http://www.elgato.com/elgato/int/mainmenu/products/Turbo264HD/product1.de.html> - letzter Zugriff am 06.03.2013.

<sup>3</sup> <http://www.videohelp.com/tools/TSSnipper> - letzter Zugriff am 06.03.2013.

<sup>4</sup> [http://www.any-video-converter.com/products/for\\_video\\_free](http://www.any-video-converter.com/products/for_video_free) - letzter Zugriff am 06.03.2013.

stellte besonders der Semesterbeginn eine Herausforderung dar. Die oftmals späte Bekanntgabe beziehungsweise manchmal nachträgliche Änderung der Kurse, welche die Studierenden im kommenden Semester besuchen werden, waren ein großes Problem für den rechtzeitigen Beginn der Aufzeichnungen und die Organisation der benötigten Tutoren/-innen.

Eine weitere Schwierigkeit stellte auch die späte Zustimmung mancher Vortragender dar. Professoren/-innen reagierten oft sehr spät oder gar nicht auf die E-Mail Anfragen und waren manchmal nur schwer telefonisch zu erreichen. Dabei kam es auch häufig zu Ablehnungen durch die Lehrenden. In vielen Fällen betraf das die Verwendung der Aufnahmen zum Training der EML ASR. Insgesamt lehnte etwa die Hälfte eine Aufzeichnung ihrer Vorlesung, trotz der Option diese nur den hörbehinderten Studierenden zur Verfügung zu stellen, ab. Das wurde meistens damit begründet, der Vortrag beziehungsweise die verwendeten Unterlagen seien nicht für eine Aufnahme auf Video konzipiert. Ein weiterer Grund war die Sorge, die Studierenden würden die Vorlesung nicht mehr besuchen und sich lediglich die Videos ansehen. Zuletzt nannten manche Vortragende, dass die Aufzeichnung für sie zu viel zusätzlichen Stress oder Aufwand verursachen würden.

Das Argument der sinkenden Besuchszahlen wurde bereits in zahlreichen Studien entkräftet beziehungsweise relativiert, wie bereits im ersten Teil dieses Kapitels 5.1.5 auf Seite 51 gezeigt wurde. Die im Zuge des GESTU Projekts durchgeführten Befragungen der teilnehmenden hörbehinderten Studenten/-innen ergaben, dass sie trotz des Vorhandenseins der Aufzeichnungen die Vorlesungen sehr oft besuchten. Die Bewertung durch die hörbehinderten Studierenden ist im nächsten Abschnitt auf Seite 68 beschrieben.

Insgesamt gab es bei den meisten Lehrveranstaltungen, die aufgezeichnet wurden, ein großes Entgegenkommen der Vortragenden. Manche Professoren/-innen äußerten die Frage, ob der Aufwand für nur eine/-n Studierende/-n gerechtfertigt sei. Beim GESTU Projekt und bei den dabei erprobten technischen Maßnahmen handelt es sich nicht um die Bevorzugung der hörbehinderten Studierenden, sondern um die Schaffung von Chancengleichheit im Bildungssystem (siehe [KS07, S: 369 - 370], [Lar10]). Auch der Nutzen der Lehrveranstaltungsaufzeichnungen wird durch die Aussagen im Zuge der Befragungen der Teilnehmer/-innen des Modellversuchs unterstrichen.

Die Aufnahmen sind auch auf großes Interesse aller Besucher/-innen der Vorlesungen gestoßen. Wurden sie allen am Kurs teilnehmenden Studierenden zugänglich gemacht, konnte damit nicht nur ein Mehrwert für die Hörbeeinträchtigten, sondern für alle Studenten/-innen geschaffen werden. Das zeigte sich auch durch das mündliche Feedback während der einzelnen Aufzeichnungen.

Besonders positiv gestaltete sich auch die Zusammenarbeit mit den, mit der Lehrveranstaltungsaufzeichnung an den einzelnen Universitäten beauftragten Abteilungen. Dadurch konnten die Aufnahmen nicht nur kostengünstig und einfach durchgeführt werden, sondern auch nahtlos in das bereits vorhandene Angebot eingefügt werden.

### **Synote**

Während des GESTU Projekts wurden auch immer wieder Versuche mit der zuvor beschriebenen E-Learning-Plattform Synote durchgeführt (siehe Abschnitt 5.2 auf Seite 57). Wie bereits eingangs in diesem Abschnitt sowie in der Motivation für den Einsatz der EML ASR im Kapitel

3.2 auf Seite 24 beschrieben, war es ursprünglich ein Ziel, die Aufzeichnungen gemeinsam mit den durch das Spracherkennungssystem erstellten Transkripten und den Präsentationsfolien in Synote zu verknüpfen. Durch die ungenügenden Resultate der Spracherkennungssoftware konnte das jedoch während der Laufzeit des Modellversuchs nie ausprobiert werden.

Im letzten Semester des Projekts fanden dennoch einige Versuche mit der E-Learning-Plattform statt. Dabei wurden Aufzeichnungen gemeinsam mit Transkripten, die durch die Respeaking Technik erstellt wurden, und den Präsentationsfolien in Synote eingebunden. Vor allem die Beschaffung und Einbindung der Folien stellte dabei ein großes Problem dar. Der betreffende Vortragende änderte sie oftmals kurzfristig und leitete sie auch nur mit sehr großer Verzögerung weiter. Die Einbindung in die E-Learning-Umgebung nahm sehr viel Zeit in Anspruch.

Grundsätzlich zeigte sich bei den durchgeführten Versuchen, dass der Einsatz von Synote an den Wiener Universitäten und den von ihnen eingesetzten Streaming-Infrastrukturen möglich ist. Dabei handelte es sich jedoch lediglich um vereinzelte Experimente. Deshalb kann keine allgemeine Aussage über den dafür benötigten Zeitaufwand getroffen werden. Durch die Versuche zeigte sich, dass es durch die Verwendung von Synote möglich ist, die technischen Limitierungen der, an den Wiener Universitäten eingesetzten Streaming-Infrastrukturen zu umgehen. Somit können auch Untertitel eingesetzt werden, wo das zuvor technisch gar nicht möglich wäre.

In mündlich durchgeführten Befragungen beurteilte die hörbehinderte Studentin den Einsatz von Synote als sehr positiv. Das deckt sich auch mit den Ergebnissen von Wald (siehe [Wal10]). Aufgrund der geringen Anzahl an Einsätzen ist es schwierig, eine allgemeine Bewertung für die Verwendung der Plattform für hörbehinderte Studierende an den Wiener Universitäten abzugeben. Weiters fehlte während der Laufzeit des GESTU Projekts eine effektive und kostengünstige Möglichkeit zur Erstellung der Transkripte. Hierbei könnte jedoch der Einsatz der Respeaking Technik und des von Hattinger im Zuge seiner zweiten Diplomarbeit [Hat13] erarbeiteten Lehrgangs, wie auf Seite 27 in Abschnitt 3.2.3 beschrieben, eine mögliche Lösung bieten.

Somit sind für den Alltagsbetrieb von Synote an den Wiener Universitäten noch weitere Versuche notwendig. Die reibungslose Planung und Durchführung der Lehrveranstaltungsaufzeichnung ist jedoch bereits ein wesentlicher Schritt zu einem möglichen späteren Einsatz der E-Learning-Plattform.

### **Bewertung durch die hörbehinderten Studierenden**

Bei jedem, im GESTU Projekt eingesetzten technische Hilfsmittel, ist die Annahme und Bewertung durch die hörbehinderten Studierenden von entscheidender Bedeutung zur Beurteilung einer Maßnahme. Das trifft besonders auf die Lehrveranstaltungsaufzeichnungen zu. Deren direkter Beitrag zum Abbau von bestehenden Barrieren ist oftmals nicht sofort ersichtlich und wird dadurch gerne als mögliche Zusatzoption ohne großen Nutzen abgetan. Bei der Befragung der hörbehinderten Studenten/-innen, die im Laufe des Modellversuchs die Lehrveranstaltungsaufzeichnungen verwendeten, stand daher die Verwendung der Maßnahme im Mittelpunkt. Weiters sollte auch ergründet werden, welche positiven Auswirkungen die Anwendung mit sich bringt, ob es dadurch möglich war, bestehende Hürden abzubauen und den Studienalltag für die Studierenden zu erleichtern.

Insgesamt erprobten sechs Teilnehmer/-innen des Modellversuchs die Vorlesungsaufzeichnungen. Vier davon wurden im letzten Projektjahr über ihre Erfahrungen zum Einsatz in sieben

verschiedenen Kursen befragt. Es handelte sich um Fächer der Studienrichtungen Wirtschaftsinformatik, Psychologie und Literaturwissenschaft. Die Befragung wurde für jede Lehrveranstaltung einzeln durchgeführt. Es wurden insgesamt sieben Nennungen verzeichnet:

- Von den sieben Kursen wurden bei fünf „75% oder mehr“ und bei zwei „50 - 75%“ aller Termine besucht.
- Bei sechs der sieben Kurse wurde die Lehrveranstaltungsaufzeichnung als nützlich bewertet.  
Bei der einen negativen Nennung gab der Student an, dass der Professor in der Vorlesung sehr viel an der Tafel schrieb. Das Tafelbild wurde jedoch nicht aufgezeichnet.
- Bei vier Kursen begünstigte die Aufzeichnung den Abschluss der Lehrveranstaltung.  
Bei den drei Nennungen, die besagten, dass die Aufzeichnungen den Abschluss der Vorlesung nicht begünstigen, wurde einmal angegeben, dass der Kurs keine Prüfung als Grundlage für die Leistungsbeurteilung hatte. Daher waren die Aufnahmen für den Abschluss nicht maßgeblich.  
Bei einer Nennung war zum Zeitpunkt der Befragung die Prüfung noch nicht absolviert. Die dritte Nennung war jene Vorlesung, bei der der Professor sehr viel an der Tafel schrieb.

Zur Nutzung der Aufzeichnungen gaben die Studierenden folgendes an (hierbei waren Mehrfachnennungen erlaubt):

- Die Studenten/-innen sahen sich bei drei Kursen die gesamte Aufzeichnung noch einmal an.  
Bei fünf Kursen sahen sich die Studierenden gezielt einzelne Abschnitte nochmals an.
- Dabei gaben die Studierenden an, dass sie die Aufnahmen zur Nachbearbeitung des Stoffes verwendeten. Die Aufzeichnungen wurden dafür als wesentlich bezeichnet.

Bei den Lehrveranstaltungsaufzeichnungen bewerteten die Studenten/-innen die Verständlichkeit der ÖGS Dolmetscher/-innen durchwegs als gut bis sehr gut. Die schlechteren Bewertungen bei der Frage hatten jedoch in allen Fällen keine technischen Gründe. Hierbei wurde zum Beispiel die Müdigkeit der Dolmetscherin genannt. Auch die Videoqualität wurde als gut bis sehr gut bezeichnet. Bei der Priorisierung der einzelnen Informationsquellen der Vorlesung wurde nachfolgende Reihenfolge angegeben:

1. Die ÖGS Dolmetscher/-innen.
2. Die Präsentationsfolien gleichwertig mit dem Tafelbild.
3. Die vortragende Person.

Zum Abschluss der Befragung hatten die Studierenden die Gelegenheit, die Vorteile der Lehrveranstaltungsaufzeichnungen zu nennen:

- Die Verfügbarkeit der Videos im Internet.
- Die schwarze Kleidung der Dolmetscher/-innen und der daraus resultierende gute Kontrast zu deren Händen.
- Die Bildqualität der einzelnen Videos.
- Der Einsatz von zwei Videokameras zur Aufzeichnung der Gebärdensprachdolmetscher/-innen und der Diskussion der gehörlosen Studierenden.
- Die Benachrichtigung über die Verfügbarkeit des Videos.

Folgende Punkte wurden als negativ bewertet:

- Beide Streaming-Server verwendeten zur Wiedergabe Adobe Flash. Daher konnten die Videos nicht auf mobilen Geräten der Firma Apple - zum Beispiel dem iPad - wiedergegeben werden.
- Die Dauer von zehn Tagen, bis die Videos verfügbar waren, wurde als zu lange bewertet.<sup>5</sup>
- Das Video der Aufzeichnung einer Lehrveranstaltung war zu dunkel.<sup>6</sup>
- Trotz der Aufzeichnung im 16:9 Bildformat wurde neben den Gebärdensprachdolmetschern/-innen weder die Tafel noch die Folienpräsentation mit aufgezeichnet.

Die Antworten und die Ergebnisse der Befragung bestätigen die im Abschnitt „Lehrveranstaltungsaufzeichnungen für hörbehinderte Studierende“ auf Seite 54 zusammengefassten Vorteile von Lehrveranstaltungsaufzeichnungen für hörbehinderte Studierende. Es wird vor allem die Möglichkeit der Nachbearbeitung des Inhaltes der Vorlesung durch die Aufnahmen als sehr wichtig bewertet. Sie gewinnt bei ungünstigen Bedingungen während der Vorlesungen, wie zum Beispiel einer hohen Schwierigkeit des Vortrages oder wenn der Stoff sehr schnell präsentiert wird, stark an Bedeutung. Das zur Verfügung stehen der Aufzeichnungen kann auch manche Probleme, die mit dem Einsatz der Gebärdensprache zur Übersetzung des Inhaltes von Vorlesungen für hörende Menschen entstehen, entschärfen. Die nachfolgenden Aussage eines am Projekt teilnehmenden gehörlosen Studierenden streicht das hervor:

- „Im österreichischen Bildungssystem werden die gesprochenen Vorträge nahezu ausschließlich mithilfe von ÖGS Dolmetschern/-innen für gehörlose Studierende übersetzt und somit zugänglich gemacht. Oftmals wird in den Vorlesungen unterstützendes Bildmaterial gezeigt. Da es sich bei Gebärdensprachen um visuelle Sprachen handelt, ist es sehr schwierig, der Übersetzung der Dolmetscher/-innen zu folgen und gleichzeitig die gezeigten Materialien zu betrachten.“

<sup>5</sup> Das war bei zwei Lehrveranstaltungsaufzeichnungen der Fall.

<sup>6</sup> Anmerkung des Autors: der Vortragende verdunkelte in der Vorlesung den Hörsaal zum Teil sehr stark.

- „Der Vortrag eines Professors ist didaktisch für hörende Studierende konzipiert. Bei der Übersetzung in die Gebärdensprache orientieren sich die Dolmetscher/-innen natürlich an der Geschwindigkeit sowie den verwendeten Ausdrücken und Lehrmethoden der Vorlesung. Diese sind für gehörlose Studierende aufgrund des sprachkulturellen Unterschiedes oftmals nicht so leicht zu verstehen.“<sup>7</sup>
- „Zuletzt halfen die Lehrveranstaltungsaufzeichnungen auch beim Wiederholen des Stoffes. Wenn zum Beispiel Sachverhalte aus den vorhandenen Unterlagen nicht verständlich waren, boten die Aufzeichnungen die Möglichkeit den Inhalt der Vorlesung nochmals anzusehen. Erst da war es möglich dem Bildmaterial im Unterricht zu folgen, da bei Aufzeichnungen die Wiedergabe pausiert werden konnte.“

„Aus diesen Gründen waren Videoaufnahmen der ÖGS Dolmetscher/-innen sprachkulturell eine absolute Notwendigkeit für mich, um wirklich barrierefrei studieren zu können.“

### 5.3.3 Empfehlungen

Bei der Evaluierung der Lehrveranstaltungsaufzeichnungen zeigte sich das große Potenzial dieses technischen Hilfsmittels. Auch wenn die Vorlesungsinhalte für die hörbehinderten Studierenden dadurch nicht direkt zugänglich gemacht werden können, wie das zum Beispiel durch ÖGS Dolmetschern/-innen geschieht, so können mit dem Einsatz von Aufzeichnungen dennoch Barrieren abgebaut werden. Das geht durch die Rückmeldung der Teilnehmer/-innen des GESTU Projekts, aber besonders durch das zuvor genannte Zitat hervor.

Wie bei jedem technischen Hilfsmittel muss auf die Rahmenbedingungen in den Kursen und besonders auf die Bedürfnisse der einzelnen Personen geachtet werden.

Aufgrund der zuvor genannten Gründe ist der weitere Einsatz von Lehrveranstaltungsaufzeichnungen als technisches Hilfsmittel zur Erleichterung des Studienalltages für hörbehinderte Studierende naheliegend und empfehlenswert. Es kann somit einen Beitrag zum Abbau bestehender Barrieren leisten. In diesem Zusammenhang ist auch die Verwendung der E-Learning-Plattform Synote zur besseren Zugänglichkeit der Inhalte der Aufnahmen erstrebenswert. Dafür sind, wie bereits zuvor beschrieben, noch weitere Versuche und eine kostengünstige Möglichkeit zur Erstellung von Transkripten erforderlich, um sie reibungslos und effektiv in den Alltagsbetrieb einbinden zu können.

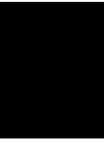
---

<sup>7</sup> Anmerkung des Autors: diese Tatsache wird auch bereits durch wissenschaftliche Studien nahegelegt (siehe [MPC<sup>+</sup>05]).



KAPITEL

6



# Fachgebärdenlexika

## 6.1 Stand der Technik und Forschung

### 6.1.1 Problem fehlender Gebärden für lautsprachliche Ausdrücke

Gebärdensprachen wurden lange nicht als vollwertige Sprachen anerkannt, da sie unter anderem als einfaches System von Pantomimen angesehen wurden. Die Österreichische Gebärdensprache (ÖGS) wurde erst im Jahr 2005 als Minderheitensprache anerkannt. Die Situation war nicht nur hierzulande sondern in vielen anderen Ländern ähnlich. Die Forschung auf dem Gebiet der Gebärdensprachen begann erst Mitte des 20. Jahrhundert - also vergleichsweise sehr spät. Deswegen fehlen beim Einsatz dieser Sprachen in vielen Bereichen wie zum Beispiel an Schulen, Universitäten, u.a.m. nach wie vor Gebärden für viele - zum Teil fundamentale - Ausdrücke. Im Bildungswesen wurde diese Situation durch die (inzwischen revidierten) Beschlüsse des Mailänder Kongresses aus dem Jahr 1880 noch verschärft. Damals wurde entschieden, dass bei der Bildung gehörloser Personen Laut- den Gebärdensprachen vorzuziehen sind. Das betrifft den tertiären Bildungssektor außerordentlich stark, denn dort fehlt es sehr oft an Gebärden für grundlegende und besonders Fachbegriffe in den einzelnen wissenschaftlichen Disziplinen. Das ist beim Einsatz der ÖGS an den Universitäten zum Teil ein großes Problem. Dadurch werden Fachausdrücke nicht, falsch oder zu einfach übersetzt (siehe [Hat11, S: 24 - 28], [Zwi10]). Krausneker und Schalber fassen dies im folgenden Zitat zusammen:

„Jones (2006) weist auch auf das Problem hin, dass an den Universitäten Fachbegriffe oder Entlehnungen oft aus dem Englischen sind und Studierende häufig die vollständige Bedeutung dieser Fachausdrücke nicht kennen oder sie (aus Mangel an Fachgebärden) zu einfach übersetzt werden, ohne die gesamte Bedeutung widerzuspiegeln. Ein adäquates Dolmetschen der Lehrinhalte wird den GebärdensprachdolmetscherInnen dadurch erschwert. In der Praxis wird dieser Mangel mit verschiedenen Strategien kompensiert (z.B. durch Buchstabieren von Fachausdrücken, was jedoch sehr zeitaufwändig ist und wobei auch Inhalte verloren gehen können, durch das Übernehmen von Gebärden anderer Gebärdensprachen, oder durch ein Übereinkommen zwischen DolmetscherIn und gehörloser Person über eine bestimmte Gebärde für einen bestimmten Fachausdruck). Das bedeutet wiederum, dass sich Studierende und DolmetscherInnen separat treffen müssen, um das Fachvokabular zu besprechen und sich auf eine Gebärde zu einigen, was einen zusätzlichen Zeitaufwand bedeutet (Interview 55). Diese Strategien können aber nicht darüber hinwegtäuschen, dass es einen großen Bedarf an Fachwörterbüchern/multimedialen Wörterbüchern zu den verschiedenen Studienbereichen gibt, wie z.B. Psychologie, Mathematik, Sprachwissenschaften oder Biologie etc.“ [KS07, S: 353].

Wie im oben angeführten Zitat beschrieben, einigen sich in manchen Fällen die beauftragten Dolmetscher/-innen und gehörlosen Studierenden vor den Vorlesungen auf Gebärden für Ausdrücke, wo diese fehlen. Sie geraten jedoch in der Regel spätestens nach dem Semester, oft schon nach der jeweiligen Vorlesungseinheit wieder in Vergessenheit, weil sie nicht festgehalten werden. Darunter leiden nicht nur die aktuell davon Betroffenen sondern auch nachfolgende Generationen gehörloser Studenten/-innen aber auch die Österreichische Gebärdensprache als Ganzes. Es ist ein wichtiges Anliegen, Möglichkeiten zu schaffen, mit denen fehlende Gebär-

den gesammelt werden können. In anderen Ländern, in denen das Problem ebenfalls besteht - zum Beispiel in Deutschland oder den USA - gibt es bereits spezielle Lexika und Internet-Plattformen zur Dokumentation von Gebärden für Fachausdrücke (siehe [BOD<sup>+</sup>08], [Cav10], [KS07, S: 353]).

Bei der Erstellung von Gebärdensprachlexika und solcher Internet-Plattformen sind einige Fakten zu beachten, die speziell im Zusammenhang mit der Natur von Gebärdensprachen auftreten. Sie werden im nächsten Abschnitt behandelt.

### 6.1.2 Herausforderungen bei der Erstellung von Gebärdensprachlexika

Lexika für Gebärdensprachen zu erstellen, ist kein leichtes Unterfangen. Es sind einige Schwierigkeiten zu lösen, die aufgrund der einzigartigen Natur dieser Sprachen und den zum Teil großen Unterschieden zu Lautsprachen existieren. Die Probleme werden besonders im Hinblick auf die Erstellung von Internet-Plattformen zur Sammlung von fehlenden Gebärden für Fachbegriffe in den nachfolgenden Abschnitten zusammengefasst (siehe [Zwi10]).

#### Darstellung der Gebärden

Die erste Frage bei der Erstellung eines Gebärdensprachlexikons ist die Darstellung der Gebärden. Dafür gibt es besser und schlechter geeignete Möglichkeiten. Bei Gebärden handelt es sich um eine manuelle Tätigkeit, die aus Gestik und sehr oft auch Mimik besteht und bei der die zeitliche Abfolge eine entscheidende Rolle spielt. Sie können daher nicht auf einfache Weise niedergeschrieben werden. Somit existieren verschiedene Formen und Möglichkeiten zu deren Darstellung (siehe [Zwi10]):

- **Bilder und Zeichnungen**

Die ersten Gebärdensprachlexika verwendeten Bilder oder Zeichnungen zur Darstellung der Gebärden. Das ist heute auch bei gedruckten Versionen noch oft der Fall. Dabei ergeben sich jedoch einige Probleme. Mit Abbildungen können nur schwer Bewegungen und Veränderungen der Mimik erfasst werden. Zur Kompensation werden sehr oft Pfeile und andere Symbole eingefügt. Da Bilder sehr viel Platz benötigen, gibt es kaum Lexika, die sie zur Darstellung der Gebärden verwenden und gleichzeitig eine große Anzahl an Einträgen beinhalten. Aber auch das Hinzufügen weiterer Informationen zu den einzelnen Einträgen in der Gebärdensprache ist kaum möglich. Das wird daher meist durch die Verwendung von Texten gelöst. Bilder sind nur schwer zu durchsuchen und zu sortieren. Daher wird meist auf eine alphabetische Gliederung der lautsprachlichen Übersetzung zurückgegriffen (siehe [Zwi10]).

In der Abbildung 6.1 ist der Eintrag für „Bär“ eines ASL Lexikons zu sehen und in Abbildung 6.2 die ÖGS Gebärde für „GESTU“.

- **Notationssysteme**

Die Niederschrift von Gebärden durch Symbole und Zeichen ist eine weitere Form der Darstellung. Das ist jedoch kein leichtes Unterfangen. Im Gegensatz zur gesprochenen Sprache handelt es sich um keine Abfolge von Buchstaben, sondern um komplexe räumliche Bewegungen, wobei oft auch die Mimik entscheidend ist. Notationssysteme für Ge-



**Bear** (1) *n* a large, heavy mammal with thick, rough fur: *a grizzly bear*.

[Action of a bear scratching itself]

With the arms crossed at the wrist on the chest, scratch the fingers of both *curved hands* up and down near each shoulder with a repeated movement.

**Abbildung 6.1:** Der Eintrag für „Bär“ in einem ASL Lexikon [Zwi10, Figure 2].



**Abbildung 6.2:** Die ÖGS Gebärde für „GESTU“, dargestellt durch zwei Bilder [Hat11, Abbildung 1.2].

|  |                                     |
|--|-------------------------------------|
| Stokoe notation system<br>(Stokoe et al. 1965) | [ ] √C <sup>†</sup> √C <sup>x</sup> |
| HamNoSys                                       | .. 𐄎 r 0 X ■ 𐄎 X [ ↓ → 𐄎 ] +        |
| SignWriting<br>(DAC 2010)                      |                                     |

**Abbildung 6.3:** Die Darstellung der ASL Gebärde für „Bär“ in drei verschiedenen Notationssystemen [Zwi10, Figure 4].

bärden sind daher sehr komplex. Sie werden daher zumeist nur in wissenschaftlichen Kreisen, aber allgemein selten eingesetzt (siehe [Cav10, S: 72], [Hat11, S: 27]), [Zwi10]).

In Abbildung 6.3 ist die ASL Gebärde für „Bär“ in drei verschiedenen Notationssystemen dargestellt.

- **Videos**

Im Gegensatz zu den beiden, zuvor genannten Möglichkeiten der Darstellung - besonders zu Bildern - erlauben Videos eine wesentlich bessere und leichter verständliche Darstellung von Gebärden und im speziellen der eingesetzten Mimik. Die Verwendung ist jedoch nicht nur mit Vorteilen verbunden. Werden sie professionell erstellt und in großem Umfang benötigt, sind sie aufwändig, zeitintensiv und teuer in der Produktion. Müssen die Aufnahmen keinen hohen Standards entsprechen, ist es heute mit handelsüblicher und kostengünstiger Computerhardware oder besser noch Videokameras bereits möglich, sie schnell in einer guten Qualität zu erstellen.

Ein gravierender Nachteil ist, dass die gebärdende Person klar zu erkennen und identifizieren ist. Sie muss daher der Veröffentlichung zustimmen. Bei der Form der Darstellung können auch computergestützte Personen eingesetzt werden. Die sogenannten Avatare bieten eine anonyme Form der Darstellung, wirken aber künstlich. Zu deren Erstellung und Anzeige muss ein geeignetes System entwickelt oder gefunden werden. Selbst damit ist die Produktion einzelner Gebärden meistens sehr aufwändig. Ist das erst einmal geschehen, können sie relativ leicht wieder geändert werden, ohne das gesamte Video neu aufnehmen zu müssen. Weiters können bei der Verwendung von Avataren die Rahmenbedingungen ihrer Darstellung besser kontrolliert werden. Das betrifft zum Beispiel das Aussehen der künstlichen Person, den Hintergrund und die Beleuchtungsbedingungen. Sie benötigen im Vergleich zu herkömmlichen Videos weniger Speicherplatz (siehe [Hat11, S: 80 - 84], [Zwi10]).

### **Art des Lexikons**

Lexika sind in der Regel Nachschlagewerke mit Informationen zu den jeweiligen Einträgen. Es kann sich unter anderem um die korrekte Trennungsweise, Pluralbildung, Fälle, eine Erklärung oder die Zuordnung zu einem Ausdruck in einer anderen Sprache handeln. Ist ein Lexikon in nur einer Sprache gehalten, wird es als monolingual bezeichnet. Enthält es zwei Sprachen ist es ein bilinguales Lexikon.

In der Vergangenheit handelte es sich bei „Gebärdensprachlexika“ meist um bilinguale Wortlisten. Sie ordneten eine Gebärde einem lautsprachlichen Wort zu. Das geschah vor allem aufgrund der fehlenden Anerkennung von Gebärdensprachen aber auch aus der Verwendung von Bildern zur Darstellung von Gebärden (siehe [Zwi10]).

### **Übersetzung von Gebärden**

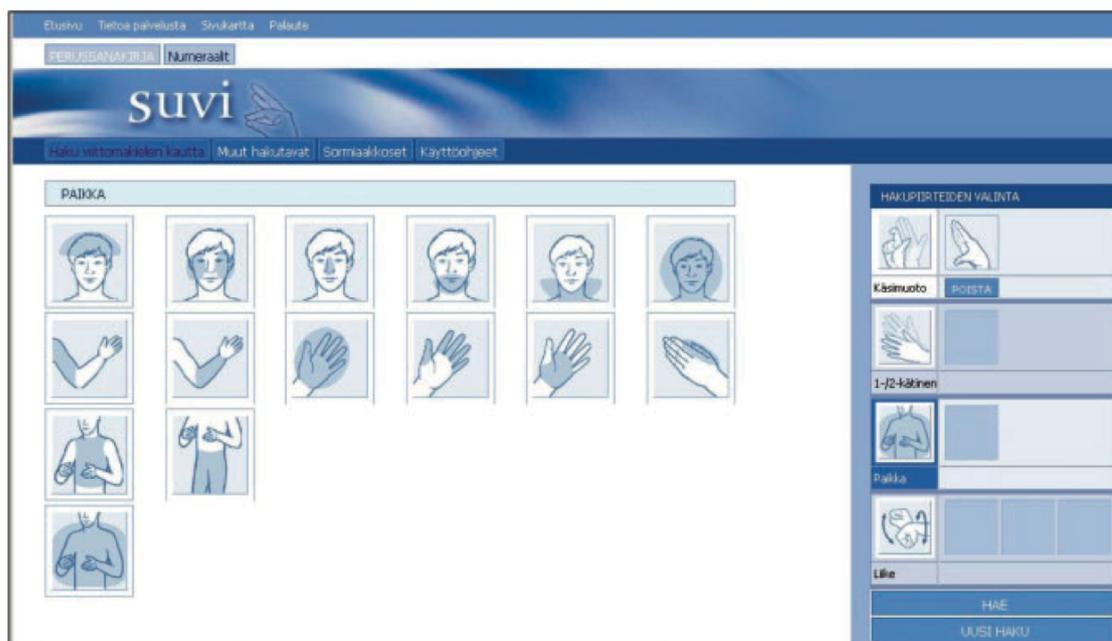
Gebärdensprachen und einzelne Gebärden unterscheiden sich zum Teil sehr von gesprochenen Sprachen und deren Wörtern. Eine Gebärde stellt oftmals ganze Tätigkeiten oder Sachverhalte dar, die lautsprachlich oft nur von ganzen Sätzen beschrieben werden können. Umgekehrt kann

das jedoch auch der Fall sein. Diese Tatsache ist bei der Erstellung eines bilingualen Lexikons ein großes Problem. Schließlich soll eine möglichst einfache - idealerweise mit einem Wort - aber dennoch korrekte Übersetzung erfolgen (siehe [Zwi10]).

## Suche

Die Suche in Lexika kann nur in eine Richtung - also unidirektional - möglich sein. Bei der Art der Suche können zum Beispiel Gebärden nur anhand von geschriebenen Wörtern gesucht werden. Die zweite Variante ist die bidirektionale Suche. Dabei ist es möglich, sowohl anhand von Wörtern nach der Bedeutung von Gebärden zu suchen aber auch umgekehrt nach Gebärden durch Angabe gewisser Merkmale. Es handelt sich dabei um die Handform, den Ausführungsort, die Bewegung während des Gebärdens oder die Mimik. Das kann entweder durch Gebärdennotationssysteme oder Bilddarstellungen erfolgen. Die Suchmaske eines Online-Gebärdenlexikons zur Suche von Gebärden anhand von Gebärdenmerkmalen ist in Abbildung 6.4 zu sehen.

Die bidirektionale Suche wurde erst durch das Aufkommen moderner Computertechnik ermöglicht. Dennoch ist die Erstellung solcher Lexika, die eine Suche in beide Richtungen erlauben, sehr aufwändig. Aus diesem Grund gibt es immer noch viele unidirektionale Gebärdensprachlexika (siehe [Zwi10]).



**Abbildung 6.4:** Suchmaske zur Auswahl des Ausführungsortes der Gebärde in einem online Gebärdenlexikon [Zwi10, Figure 8].

### 6.1.3 Fachgebärdensammlung

Zu Beginn des Kapitels wurden die Probleme von Gebärdensprachen durch mangelnde Gebärden für bestimmte Wörter und die Schwierigkeiten bei ihrer Verwendung im tertiären Bildungssektor beschrieben. Nachdem auch kurz die Grundlagen und Herausforderungen bei der Erstellung von Gebärdensprachlexika erörtert wurden, ist zu klären, wie überhaupt Gebärden „erfunden“ beziehungsweise in den allgemeinen Gebrauch gelangen. Dazu gibt es verschiedene Ansätze, von denen einer die Einsetzung einer Kommission oder Gruppe von Experten/-innen zur Sammlung und Dokumentation ist. Bei dem Prozess werden die Gebärden nach bestimmten - oftmals auch anhand linguistischer - Kriterien ausgewählt. Bis auf diesem Weg neue Gebärden beschlossen werden, dauert es in der Regel sehr lange. In der Vergangenheit traf sich ein solches Gremium für die ÖGS auf regelmäßiger Basis. Aktuell finden nach meinem Wissensstand keine Treffen statt.

Sprachen entwickeln sich auch durch ihren Gebrauch im Alltag selbstständig weiter. Neue Ausdrücke oder Gebärden entstehen, andere werden aus anderen Sprachen übernommen und manche verschwinden fast gänzlich. Im Gegensatz zu den Lautsprachen, die durch jahrtausendelange Niederschrift sehr gut dokumentiert sind, fehlte es bei Gebärdensprachen lange an guten und einfachen Möglichkeiten zum Festhalten einzelner Gebärden.

Dabei entstehen auch bei der Verwendung von Gebärdensprachen im tertiären Bildungssektor und in der Wissenschaft ständig neue Gebärden. Es werden von den Studierenden, Dolmetschern/-innen und von in diesen Gebieten arbeitenden gehörlosen Personen ständig neue Fachgebärden entwickelt und eingesetzt:

„Today, deaf scientists, students, and professionals often decide on signs to use on a local (regional) or temporal (the duration of a class or conference) basis. As a result, alternative signs for the same term are developed, and signs that are developed are lost“ [Cav10, S: 72].

Sie werden jedoch nicht oder nur selten dokumentiert, da sie nach einer Lehrveranstaltung oder einem Treffen meist nicht mehr benötigt werden. Eine weitere Möglichkeit zur Findung, Sammlung und Etablierung von Fachgebärden ist es, den ständigen Prozess von Neuerfindung, Verwendung und Vergessen zu unterbrechen und sie zu dokumentieren. Die Vorschläge für neue Gebärden kommen bei dem Ansatz von Personen, welche die Gebärdensprache in den jeweiligen (wissenschaftlichen) Fachgebieten verwenden. Durch den täglichen Sprachgebrauch sind sie darin Experten/-innen. Dazu muss ihnen jedoch eine gute Möglichkeit eröffnet werden, die Gebärden zu sammeln, zu diskutieren, zu bewerten und so eine gute Variante auszuwählen. Das ist heutzutage unter Verwendung moderner Computer- und Videotechnik sowie dem Internet möglich. Rector et al. beschreibt dazu die Verwendung einer speziellen Multimedia-Plattform im Internet - eines Forums:

„A forum is a dynamic growing version of a dictionary, meaning signs can be entered based on people’s needs and there can never be too much information. A dictionary moves at a much slower pace, and if a popular term is introduced there can be delays of up to years before it is introduced [2]. Unlike a dictionary in

which employers are experts in the language, contributions in a forum come from the users“ [RLS11].

Die so gesammelten Gebärden können somit sehr schnell eingesetzt und eventuell später in ein offizielles Gebärdenlexikon aufgenommen werden. Cavander zeigt in ihrer Dissertation, dass das inzwischen sehr oft bei gesprochenen Sprachen üblich ist. Ihre Untersuchung ergab, dass von den 657 Begriffen, die 2008 in das Oxford English Dictionary neu aufgenommen wurden, 450 bereits zuvor bei Wikipedia<sup>1</sup>, Wiktionary<sup>2</sup> oder dem Urban Dictionary<sup>3</sup> zu finden waren (siehe [Cav10, S: 73 - 74]).

#### 6.1.4 ASL-STEM Forum

Beim ASL-STEM Forum<sup>4</sup> handelt es sich um eine erfolgreiche und sehr gute Umsetzung der zuvor beschriebenen Möglichkeit zur Sammlung von Fachgebärden durch die Gebärdensprachbenutzer/-innen in Form einer Internet-Plattform. Der Name setzt sich aus den Wörtern „ASL“ für American Sign Language und „STEM“ - der Anfangsbuchstaben für die englischen Fachbereiche Science, Technology, Engineering und Mathematics - zusammen. Es wurde erstellt, um die eingangs beschriebenen Probleme der fehlenden Fachgebärden in der amerikanischen Gebärdensprache zu verringern:

„There have been efforts to develop and disseminate signs for STEM topics, often through careful consideration and by expert committee [50, 91, 92, 148]. These video sign language dictionaries indicate a strong need for consensus among members of the deaf community about signing for STEM topics and have been an important first step. However, natural languages evolve by consensus not by committee [125]. In all written, spoken, or signed languages, vocabulary arises from the language-using population, with standardization resulting from popular adoption over time. This type of natural language evolution is the goal of the ASL-STEM Forum“ [Cav10, S: 73].

Die Autoren geben an, dass das ASL-STEM Forum zum Sammeln und Nachschlagen von Gebärden im Internet dienen soll. Aufgrund dieser Anforderungen handelt es sich dabei um kein Lexikon:

„The ASL-STEM is not considered a dictionary; there are many differing attributes from traditional online dictionaries that are conducive to online collaboration“ [RLS11].

Nachfolgend wird die Realisierung und die Erkenntnisse durch die Verwendung näher betrachtet, um so die wichtigsten Kriterien für eine solche Internet-Plattform zu erläutern.

<sup>1</sup> <http://www.wikipedia.org> - letzter Zugriff am 06.03.2013.

<sup>2</sup> <http://www.wiktionary.org> - letzter Zugriff am 06.03.2013.

<sup>3</sup> <http://www.urbandictionary.com> - letzter Zugriff am 06.03.2013.

<sup>4</sup> <http://aslstem.cs.washington.edu> - letzter Zugriff am 06.03.2013.

### Realisierung

Im ASL-STEM Forum können unter Verwendung von Videos Gebärden für fehlende Fachausdrücke vorgeschlagen und diskutiert werden. Es bietet darüber hinaus die Möglichkeit zur Suche anhand geschriebener Wörter. Es verbindet dabei bereits im Internet übliche Prinzipien. Die Verwendung von Videos, Videokonferenzen aber auch die gemeinsame, freiwillige und unentgeltliche Zusammenarbeit zur Erstellung von allgemein zugänglichen Inhalten ist heutzutage kein ungewöhnliches Konzept mehr und hat auch bereits ihre Auswirkungen auf die Gebärdensprachen in allen Ländern:

„Our ASL-STEM Forum builds on the rapid growth of online video phones, user-contributed video sites, and video-enabled social networking sites. These video technologies have already affected ASL by bringing together dispersed deaf communities and accelerating the natural process of discussion about differing vocabulary and idioms, resulting in agreement on common vocabulary or acceptance of alternatives for the same concept“ [Cav10, S: 73].

Die Plattform beruht auf der gemeinsamen und freiwilligen Zusammenarbeit. Daher muss die Mitarbeit einfach, benutzerfreundlich, unabhängig von technischen Kenntnissen und mit geringem Aufwand möglich sein. Es sollen viele Personen zum Inhalt beitragen können und auch dazu motiviert werden. Bei den einzelnen Einträgen und für die breite Nutzung ist nicht nur die Quantität sondern auch die Qualität ausschlaggebend.

Die oben angeführten Voraussetzungen sind laut Cavender wesentlich und werden in den beiden Designkriterien des ASL-STEM Forums zusammengefasst:

- „The ASL-STEM Forum should require little technical expertise and little time commitment to contribute video signs.
- The ASL-STEM Forum should be inviting and community members should feel comfortable contributing their knowledge and opinions“ [Cav10, S: 74].

Ausgehend von den obigen Anforderungen wurde das ASL-STEM Forum unter Zuhilfenahme des Ruby on Rails Frameworks programmiert. In Abbildung 6.5 ist das Interface der Seite und der Eintrag für den Begriff „Array“ zu sehen.

Bei der Realisierung des Forums wurden die nachfolgenden Lösungsansätze gewählt:

- **Organisation der Einträge**

Ziel des Forums ist es, Gebärden ausgehend von lautsprachlichen Fachausdrücken zu sammeln. Der geschriebene Begriff wird als Name für die einzelnen Einträge und zu deren Organisation verwendet. Daher ist das ASL-STEM Forum eine bilinguale Liste von Wörtern, denen Gebärden zugeordnet sind. Aufgrund der Anforderungen und der Tatsache, dass es sich dabei um kein Lexikon im klassischen Sinn handelt, ist das jedoch ausreichend.

- **Kategorisierung der Fachbegriffe**

Das Forum selbst ist in die vier namensgebenden Kategorien „Science“, „Technology“, „Engineering“ und „Mathematics“ unterteilt. Bei Bedarf können noch weitere Unterkategorien angelegt werden. Die einzelnen Einträge sind den Kategorien zugewiesen.

The screenshot shows the ASL-STEM Forum interface. At the top, there is a navigation bar with 'Topics', 'About', and 'Help'. The forum title is 'ASL-STEM Forum' with the tagline 'Enabling American Sign Language to grow in Science, Technology, Engineering, and Mathematics (STEM)'. A login section includes fields for 'E-mail' and 'Password', and a 'Sign in' button. Below the navigation bar, the current topic is 'Arrays' under the path 'Computer Science >> Data Structures >> Arrays'. A sidebar on the left lists 'Data Structures' and 'Arrays' with sub-topics like 'Array Loops' and 'Constructing Arrays'. The main content area features a 'Topic Definition' section with a definition and an example. Below this is a 'Sign Suggestions' section with four video thumbnails, each with a '5/5' rating and an 'Enlarge' button. A large video player on the right shows a 'Highest Rated Sign' with a video ID of #91, posted on 5/19/2009 by user 'millam', with a 5-star rating. A 'Text Comments' section at the bottom shows a comment from user 'jndewitt' dated 4:08pm 3/26/2008, stating that video #91 was from a dictionary and didn't clarify arrays. The forum also includes a search bar, a 'sub-topics' button, and sorting options for comments.

Abbildung 6.5: Der Eintrag für „Array“ im ASL-STEM Forum [Cav10, Figure 5.2].

### • Aufbau eines Eintrages

Ein Fachgebärdeneintrag bietet neben dem Namen des Fachbegriffes selbst noch die Möglichkeit zur Angabe einer Definition sowie eines Beispielsatzes. Sofern ein oder mehrere Videos mit Vorschlägen für eine Gebärde existieren werden sie ebenfalls angezeigt. Der, durch die Benutzer/-innen am besten bewertete Vorschlag wird groß, die restlichen klein angezeigt. Negativ bewertete Videos werden auf einer separaten Seite angezeigt (siehe [Cav10, S: 76]). Das ist in Abbildung 6.5 dargestellt.

### • Darstellung der Gebärden durch Videos

Wie bereits im ersten Teil des Kapitels ab Seite 75 beschrieben, eignen sich Videos sehr gut zur Darstellung von Gebärden. Sie können durch heute handelsübliche Computer einfach und schnell erstellt werden. Die einzelnen Gebärdenvorschläge werden als selbst-erstellte Filme der einzelnen registrierten Benutzer/-innen dargestellt. Dadurch ist eine

einfache und unkomplizierte Mitarbeit möglich, ohne dass eines der komplexen Notationssysteme verwendet wird.

Die Videos werden auf der Internetvideoplattform YouTube<sup>5</sup> gespeichert und in die Seite eingebunden. Sie können durch ein, vom ASL-STEM Forum bereitgestelltes Applet direkt mit einer Webcam oder anderen Videokamera aufgezeichnet und anschließend eingebunden werden.

- **Kommentare**

Zu jedem Eintrag können Kommentare in schriftlicher Form erfolgen. Somit können die Inhalte und Vorschläge diskutiert werden.

- **Einfaches System zum Bewerten von Gebärden**

Neben dem Verfassen von schriftlichen Kommentaren ist es den Benutzern/-innen auch möglich, die Videos mit den Gebärdenvorschlägen mit null bis fünf Sternen zu bewerten, wobei fünf die höchste Bewertung darstellt. Anhand der Bewertungen erfolgt die Auswahl des besten Eintrages, der groß angezeigt wird. Laut Cavender ist so eine schnelle und anonyme Mitarbeit möglich:

„We think the rating system may be an attractive alternative for a specific type of user: one who does not want to setup a camera or manage video, or who would prefer a more anonymous way to still express their opinion“ [Cav10, S: 83].

- **Suche**

Zum schnellen Auffinden von Gebärden ist es möglich, die Einträge des ASL-STEM Forums und auch deren Definition durch Angabe eines geschriebenen Textes zu durchsuchen. Eine Suche anhand von Gebärdenmerkmalen wie zum Beispiel der Handform, dem Ausführungsort u.a.m. wird nicht angeboten.

- **Quantität und Qualität der Einträge**

Eines der Ziele des ASL-STEM Forums ist die Sammlung möglichst vieler Einträge. Dazu ist eine möglichst große Nutzerschaft notwendig. Dabei spielt nicht nur die Quantität sondern auch die Qualität der einzelnen Einträge eine entscheidende Rolle. Laut Cavender kommt den ersten und frühen Benutzern/-innen einer solchen Plattform eine besondere Rolle zu. Sie erstellen die ersten Inhalte und etablieren somit eine bestimmte Form von Normen und Standards. Daher wurden die ersten Einträge des ASL-STEM Forums in Zusammenarbeit mit zwei auf gehörlose Studierende spezialisierten Universitäten erstellt. Weiters ist es wichtig, die Benutzer/-innen zur Mitarbeit zu ermuntern. Um das zu erreichen, bietet das Forum die Möglichkeit, Einträge zu finden, die noch nicht vollständig sind. Dabei kann es sich zum Beispiel um fehlende Gebärdenvorschläge aber auch unvollständige Definitionen handeln (siehe [Cav10, S: 85 - 87], [RLS11]).

---

<sup>5</sup> <http://www.youtube.com> - letzter Zugriff am 06.03.2013.

### **Bewertung des ASL-STEM Forums durch gehörlose Benutzer**

Cavender führte für ihre Dissertation eine Studie über die Benutzung und Bewertung des ASL-STEM Forums durch. Dabei wurden 14 gehörlose Personen ausgewählt, die entweder studieren oder in einem der Fachbereiche tätig sind. Bei der Befragung bewerteten sie das Forum als wertvolle Ressource, die einfach zu verwenden ist. Die Möglichkeit zur Mitarbeit wurde als gut und leicht beschrieben. Im Durchschnitt benötigten die Teilnehmer/-innen zwei Minuten und 16 Sekunden zum Erstellen eines Eintrages inklusive des Videos für den Gebärdenvorschlag. Die Studie zeigte, dass Benutzer/-innen, die aktiv Gebärdenvorschläge hinzufügten, auch sonst vermehrt zu den Inhalten des Forums beitrugen (siehe [Cav10, S: 76 - 87]).

Neben den positiven Ergebnissen ist besonders eine Tatsache herausragend. Das Forum wird von den Befragten als „von“ und „für“ die Gemeinschaft der gehörlosen Personen empfunden:

„Several participants noted the importance of the Forum being “by” and “for” the deaf community. For example, “I believe the forum can help us, the deaf communities, to raise our language, which it will improve our communication ... However, this will not only benefit the deaf communities; it also will encourage the interpreters for the deaf to develop the language” “ [Cav10, S: 84].

## **6.2 Sammlung von Fachgebärden im Zuge des GESTU Projekts**

Wie eingangs in diesem Kapitel (siehe 6.1.1 auf Seite 74) beschrieben, mangelt es besonders an Gebärden für wissenschaftliche Fachausdrücke. Das stellt besonders für die Dolmetscher/-innen bei der Übersetzung in den Lehrveranstaltungen aber auch für die gehörlosen Studierenden ein großes Problem dar. Ein Ziel des GESTU Projekts war es, mit der Sammlung und Dokumentation von Fachgebärden zu beginnen. Zu dem Zweck wurde eine Kooperation mit Professor Franz Dotter, dem Leiter des Zentrums für Gebärdensprache und Hörbehindertenkommunikation an der Alpen-Adria-Universität Klagenfurt, gestartet. Als Ergebnis wurden von Professor Dotter an der Technischen Universität Wien zwei Lehrveranstaltungen mit dem Schwerpunkt der Gebärdensprachlinguistik gehalten. Darin wurde die Entwicklung von Gebärden sowie deren Beurteilung anhand linguistischer Kriterien behandelt.

Während des Modellversuchs fanden mehrere sogenannte Fachgebärdenworkshops statt. deren Ziel war, die teilnehmenden gehörlosen Studierenden aber auch Dolmetscher/-innen bei der Findung von Gebärden sowie deren Aufnahme auf Video zu unterstützen. Bereits bei den ersten Treffen zeigte sich, dass eine einfache und gute Möglichkeit zur Sammlung, Dokumentation und Diskussion fehlt. Auch Hattinger empfiehlt in seiner Diplomarbeit die Entwicklung eines solchen Fachgebärdenlexikons:

„Das Archivieren und Auffinden von (neuen) ÖGS Fachgebärden ist der primäre Zweck, den das GESTU Fachgebärdenwörterbuch erfüllen soll“ [Hat11, S: 133].

### **6.2.1 Realisierung der Plattform zur Sammlung von Fachgebärden**

Aufgrund der, im vorangehenden Abschnitt beschriebenen Probleme, wurde im Zuge des Modellversuchs in Zusammenarbeit mit Professor Dotter das sogenannte „GESTU Fachgebärden

The screenshot shows the GESTU Fachgebärden Wiki page for 'Exponent (Mathematik)'. The page layout includes a top navigation bar with options like 'Seite', 'Diskussion', 'Lesen', 'Mit Formular bearbeiten', and 'Versionsgeschichte'. A search bar is also present. The main content area is titled 'Exponent (Mathematik)' and contains a 'Bedeutung' section with a video player. The video shows a man making a hand gesture. A sidebar on the left contains a 'Navigation' menu with links like 'Startseite', 'letzte Änderungen', and 'Hilfe zum Wiki'. A table in the top right corner shows the 'Status' of the entry as 'Gebärde fertig'. Annotations in grey boxes highlight specific features: 'Navigationsmenü' (Navigation menu), 'Begriffsdefinition' (Concept definition), 'Status der Gebärde' (Status of the gesture), 'Gebärdenvorschlag' (Gesture proposal), and 'Anmerkungen' (Remarks).

Abbildung 6.6: Der Eintrag für den Begriff „Exponent“ im GESTU Fachgebärden Wiki.

Wiki<sup>6</sup> entwickelt. Es ist in Abbildung 6.6 zu sehen. Damit soll die Fachgebärdensammlung durch die Gebärdensprachbenutzer/-innen der verschiedenen wissenschaftlichen Fachgebiete nach der zuvor in Abschnitt 6.1.3 auf Seite 79 beschriebenen Methode erfolgen. Daher musste es den am Projekt teilnehmenden hörbehinderten Studierenden und Dolmetschern/-innen und allen Interessierten mit der Plattform ermöglicht werden, Gebärden für (wissenschaftliche) Fachausdrücke zu sammeln, zu diskutieren und nachzuschlagen.

Da die Erstellung der Inhalte durch die Benutzer/-innen erfolgt, wurde bei der Entwicklung nicht nur auf die technischen und linguistischen Anforderungen Wert gelegt, sondern wie auch beim ASL-STEM Forum (siehe Abschnitt 6.1.4 ab Seite 80) auf die Bedürfnisse der zukünftigen Anwender/-innen. Somit sollen viele Personen zur Mitarbeit eingeladen werden.

Aufgrund der unterschiedlichen Anforderungen einer solchen Plattform unterscheidet sich das

<sup>6</sup> <http://wiki.gestu.at> - letzter Zugriff am 06.03.2013.

GESTU Fachgebärden Wiki von klassischen online Lexika. Bei den Einträgen handelt es sich um keine offiziell anerkannten Gebärden, sondern um eine Sammlung von Vorschlägen und Arbeitsmitteln.

Bei der Entwicklung wurden die folgenden Anforderungen erarbeitet und Lösungen gewählt (die einzelnen Techniken und Begriffe wurden bereits im Abschnitt 6.1.2 ab Seite 75 dieses Kapitels beschrieben):

- **Erstellung des Inhaltes durch die Benutzer/-innen**  
Die einzelnen Einträge und Vorschläge für Gebärden des GESTU Fachgebärden Wikis werden von den Benutzern/-innen erstellt. Eine der wichtigsten Anforderung an die Plattform ist eine nicht zeitintensive und einfache Möglichkeit zur Mitarbeit.
- **Aufbau des GESTU Fachgebärden Wikis**  
Durch die Plattform sollen zu lautsprachlichen Fachbegriffen fehlende Gebärden gefunden und dokumentiert werden. Der geschriebene Facha Ausdruck dient als Name für die Einträge, zu deren Organisation und Sortierung. Die Erstellung der einzelnen Einträge soll unkompliziert und schnell möglich sein. Auf die Angabe von Merkmalen der Gebärden - wie zum Beispiel der Handform - wird verzichtet. Die Plattform ist daher als unidirektionales, bilinguales Lexikon realisiert.
- **Darstellung der Gebärden**  
Die einzelnen Gebärdenvorschläge werden durch selbsterstellte Videos der Benutzer/-innen dokumentiert. Das erfolgte wegen der guten Darstellung von Gebärden durch die Aufzeichnungen und der inzwischen einfachen und schnellen Möglichkeiten zur Aufnahme. Auf die Verwendung von Avataren wird trotz der fehlenden Anonymität der Videos verzichtet, da deren Erstellung aufwändiger und komplizierter ist. Bedingt durch die hohe Komplexität und die mangelnde Verbreitung wurde auf eine Verwendung von Gebärdennotationssystemen ebenfalls verzichtet.
- **Kategorisierung der Einträge**  
Das GESTU Fachgebärden Wiki soll zur Sammlung von Gebärden und Fachbegriffen aus den verschiedenen wissenschaftlichen Disziplinen genutzt werden. Sie sollen daher einem oder mehreren Themengebieten zugeordnet werden können. Alle Einträge einer bestimmten Kategorie sollen ebenfalls angezeigt werden können.
- **Inhalt eines Eintrages**  
Jedem Eintrag in der Plattform können ein oder mehrere Videos hinzugefügt werden. Sie sollen Vorschläge für Gebärden oder Erklärungen zur Bedeutung des Begriffes in der ÖGS beinhalten. Zusätzlich können schriftliche Definitionen und Anmerkungen gemacht werden.
- **Diskussion der einzelnen Vorschläge**  
Zu jedem Eintrag können schriftliche Kommentare und Anmerkungen erfolgen. Somit können die Benutzer/-innen die verschiedenen Vorschläge diskutieren.

- **Kennzeichnung des Entwicklungszustandes**

Jeder Eintrag ist mit einem Status gekennzeichnet. Somit kann der Entwicklungsstand für den Gebärdenvorschlag angegeben werden. Es wird dadurch ersichtlich, ob für den Fachbegriff eine Gebärde benötigt wird, oder ob bereits an einem Vorschlag gearbeitet wird, beziehungsweise ob abgestimmt werden kann und ob der aktuelle Vorschlag von den Benutzern angenommen wurde.

- **Suche von Einträgen**

Im GESTU Fachgebärden Wiki kann der schriftliche Inhalt der einzelnen Einträge durchsucht und die Ergebnisse angezeigt werden. Die Suche nach Gebärdens anhand ihrer Merkmale ist nicht möglich, da dafür keine Eingabe vorgesehen ist. Sie können anhand verschiedener Kriterien als Listen abgerufen werden, alphabetisch sortiert oder nach Themengebieten. Es können weiters alle Inhalte gesucht werden, für die noch Gebärdenvorschläge benötigt werden oder die bereits ein Video enthalten. Dadurch soll die Plattform benutzerfreundlicher sein und zur Mitarbeit einladen.

- **Zugang zur Plattform nach Anmeldung**

Bei der Entwicklung wurde festgelegt, dass die Inhalte des Fachgebärden Wikis nur für registrierte und vom GESTU Team freigeschaltete Benutzer/-innen zugänglich sind. Die Plattform sollte in der ersten Phase nur am Projekt teilnehmenden hörbehinderten Studierenden und Dolmetscher/-innen zur Verfügung stehen und erst zu einem späteren Zeitpunkt öffentlich zugänglich gemacht werden. Dennoch wurden auch Interessierte, wie zum Beispiel Teilnehmer/-innen der beiden von Professor Dotter abgehaltenen Lehrveranstaltungen, freigeschaltet.

Durch die Registrierung akzeptieren die Benutzer/-innen einige wichtige Regeln zur Verwendung der Plattform - so zum Beispiel, dass die Videos nicht öffentlich zugänglich gemacht werden dürfen.

Aufgrund der vorher angeführten Anforderungen und der gewählten Lösungen wurde zur Umsetzung der Plattform das Konzept eines Wikis gewählt. Es handelt sich um eine Webseite, in der die Benutzer/-innen die Inhalte selbst erstellen und bearbeiten können. Dafür wurde als Serverengine MediaWiki<sup>7</sup> gewählt, angepasst und um die im GESTU Fachgebärden Wiki benötigten Funktionen erweitert. Die MediaWiki Engine ist die Serversoftware, die auch bei Wikipedia verwendet wird. Ihr Einsatz bringt sowohl aus technischer Sicht als auch aus Benutzersicht zahlreiche Vorteile mit sich. Sie ist einfach zu bedienen und durch die große Bekanntheit - besonders durch Wikipedia - sind viele mit der Handhabung und dem zugrunde liegenden Konzept vertraut. Aus technischer Sicht erfüllt die MediaWiki Engine alle oben genannten Anforderungen oder sie lassen sich durch eine der zahlreichen Erweiterungen nachträglich hinzufügen. Die Verwendung ist aufgrund der Veröffentlichung als Open Source Produkt kostenlos, gleichzeitig ist sie durch den häufigen Einsatz gut erprobt und leicht zu warten.

Die Anforderungen zur Installation und zum Betrieb sind sehr gering. Die Software setzt lediglich einen Webserver - zum Beispiel Apache<sup>8</sup>, PHP<sup>9</sup> sowie eine Datenbank wie zum Beispiel

<sup>7</sup> <http://www.mediawiki.org> - letzter Zugriff am 06.03.2013.

<sup>8</sup> <http://httpd.apache.org> - letzter Zugriff am 06.03.2013.

<sup>9</sup> <http://php.net> - letzter Zugriff am 06.03.2013.

MySQL<sup>10</sup> voraus.

Bei der Entwicklung des GESTU Fachgebärden Wikis wurde beschlossen, die Videos selbst zu speichern und nicht auf eine der zahlreichen Internet-Videoplattformen - wie zum Beispiel YouTube hochzuladen. Für ein durchschnittliches Gebärdenvideo mit einer Länge von fünf Sekunden, das unter Berücksichtigung der bereitgestellten Spezifikationen (diese sind im nächsten Abschnitt zu finden) erstellt wird, wurde mit einem Speicherverbrauch von rund 250kB gerechnet. Angesichts der geringen Anforderungen konnte für die Realisierung der Plattform ein normaler Web-Server verwendet werden.

### **Empfehlungen für die Erstellung der Videos**

Die Erstellung der Videos erfolgt durch die Benutzer/-innen des GESTU Fachgebärden Wikis. Zur Unterstützung der Aufnahmen und der Nachbearbeitung wurden während der Umsetzung der Plattform sowie während der Fachgebärdenworkshops Empfehlungen zusammengestellt. Dabei wurde darauf geachtet, dass technische Laien die Aufzeichnungen durchführen werden. Somit wurde versucht, einen Kompromiss zwischen dem benötigten Aufwand und der Qualität zu finden. Die Empfehlungen stehen im Fachgebärden Wiki als Anleitung auf den Hilfeseiten oder als Leitfaden in Form eines PDF Dokuments zur Verfügung.

Zur Aufnahme der Videos können angesichts der geringen Anforderungen an die Bildqualität und der guten und leicht kontrollierbaren Rahmenbedingungen bei der Aufzeichnung alle gängigen Geräte eingesetzt werden. Dabei handelt es sich um digitale Videokameras, digitale Fotokameras, Webcams oder Mobiltelefone mit Videokamera.

Bei der Erstellung der Filme wird empfohlen, die folgenden Punkte zu berücksichtigen:

- Verwendung eines Stativs oder auflegen der Kamera.
- Blick in die Kamera.
- Zu Beginn und am Ende der Aufnahme jeweils ein bis zwei Sekunden in neutraler Position verbleiben.
- Aufnahme vor einem einfarbigem Hintergrund.
- Verwendung von einfarbiger, dunkler Bekleidung. Schwarz ist besonders gut geeignet, da ein guter Kontrast zur Hautfarbe besteht.
- Pro Video nur eine Gebärde und gegebenenfalls deren Bedeutung aufzeichnen.

Für die Einbindung der Aufnahmen in das GESTU Fachgebärden Wiki werden die folgenden Parameter zur Konvertierung der Videos empfohlen:

Als Videoformat wird das MP4 Format empfohlen. Weiters sollen sie auf eine Auflösung von 640 x 480 Pixel verkleinert werden und mit einer „Video Bitrate“ von 256 Bit/s komprimiert werden. Mit den beiden Einstellungen sind die Gebärden in den Aufzeichnungen sehr gut zu erkennen (siehe Abbildung 6.7) und sie benötigen nicht sehr viel Speicherplatz.

<sup>10</sup> <http://www.mysql.com> - letzter Zugriff am 06.03.2013.

### 6.2.2 Erkenntnisse

Die während des GESTU Projekts realisierte Plattform zur Sammlung, Dokumentation und dem Nachschlagen von Fachgebärden konnte erfolgreich, schnell, kostengünstig und als gute Lösung aus Sicht der Benutzer/-innen umgesetzt werden. Angesichts der begrenzten finanziellen Mittel des Modellversuchs wurde bereits bei der Planung und der Erstellung der Plattform kosteneffektiv gearbeitet. Das konnte vor allem durch die Verwendung der vorhandenen Ressourcen und Möglichkeiten im Rahmen des GESTU Projekts und der Technischen Universität Wien erreicht werden.

Die gesamte Planung, Umsetzung und Betreuung erfolgte sehr kostengünstig durch die vorhandenen Ressourcen des Modellversuchs, im Rahmen dieser Diplomarbeit und durch ein Praktikum an der Technischen Universität Wien. Professor Dotter begleitete außerdem alle Entwicklungsschritte und stellte sein Wissen auf dem Gebiet der Gebärdensprachlinguistik zur Verfügung. Vor allem der Einsatz und die Adaptierung der kostenlos verfügbaren MediaWiki Engine stellte eine wesentliche Erleichterung bei der schnellen Umsetzung des Vorhabens dar. Infolge der auf Seite 87 beschriebenen geringen Anforderungen durch verwendete Serversoftware aber auch wegen des geringen Speicherplatzverbrauchs der Videos wurde für den Betrieb der Plattform ein kostengünstiger Webserver angemietet.

Für die Erstellung der Inhalte und Einträge des Fachgebärden Wikis sind ebenso kaum Kosten entstanden. Sie wurden durch die Benutzer/-innen mit ihrem eigenen Equipment angefertigt. Das GESTU Projekt bot auch die Möglichkeit, die Videos mit dem durch die Lehrveranstaltungsaufzeichnungen bereits vorhandenen Equipment zu filmen. Dadurch sind auch für deren Erstellung keine zusätzlichen Kosten angefallen.

Während der Laufzeit des Modellversuchs und seit der Einführung der Plattform im November 2011 konnten im Zeitraum von etwas mehr als einem Jahr 585 Einträge gesammelt werden. Hiervon besitzen bereits 276 Vorschläge für Gebärden. Besonders im Hinblick auf die relativ geringe Anzahl von nur 32<sup>11</sup>, im System registrierten Benutzern/-innen, ist die Zahl der vorhandenen Einträge sehr groß. Die kleine Nutzerbasis lässt sich durch den nicht öffentlichen Zugang zur Plattform begründen. Es handelt sich dabei größtenteils um die Teilnehmer/-innen des GESTU Projekts, einige Dolmetscher/-innen sowie Interessierte. Begünstigt wurde die Sammlung durch die Verwendung der Plattform in den Fachgebärdenworkshops und der von Professor Dotter abgehaltenen Lehrveranstaltungen.

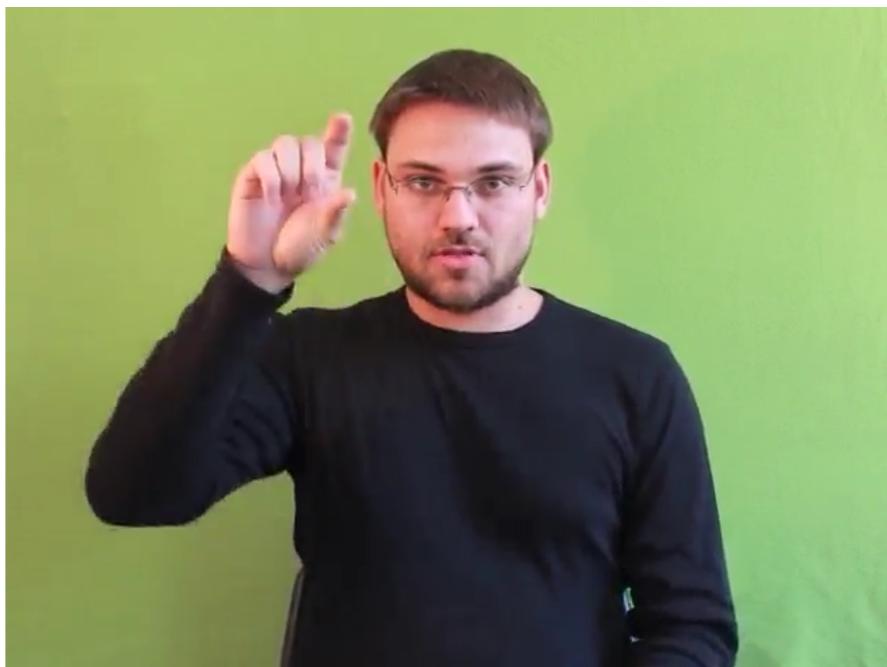
Beim Einsatz des Fachgebärden Wikis zeigte sich auch das bereits von Cavender beobachtete Verhalten, dass einige wenige Benutzer/-innen sehr viel, manche ein weniger und ein Großteil sehr wenig zu den Inhalten beitragen (siehe [Cav10, S: 84 - 85]).

Da die Benutzer/-innen die Inhalte selbst erstellen, variiert deren Qualität und Detailgrad zum Teil sehr stark. Sehr viele Einträge beinhalten lediglich den Namen des Fachausdruckes und eventuell eine kurze Erklärung. Es handelt sich meist um die Dokumentation einer fehlenden Fachgebärde. Auch die Qualität der selbsterstellten Videos schwankt je nach betriebenem Aufwand. Sie reicht von einer sehr guten Bildqualität - ein Ausschnitt aus einem solchem ist in Abbildung 6.7 zu sehen - bis hin zu einer weniger guten - die beiden Bilder in Abbildung 6.8 stammen aus zwei Videos, bei deren Erstellung nicht alle auf Seite 88 genannten Empfehlun-

---

<sup>11</sup> Die Test- und Administrationsaccounts wurden nicht berücksichtigt.

gen eingehalten wurden. Insgesamt sind jedoch alle Gebärden sowie sämtliche Details wie die Mimik oder die einzelnen Finger in den Aufnahmen gut zu erkennen.



**Abbildung 6.7:** Ein Ausschnitt aus einem, von einem Benutzer des Fachgebärden Wikis selbst-erstellten qualitativ hochwertigen Video.

Die Rückmeldungen der Benutzer/-innen des GESTU Fachgebärden Wikis ergaben, dass die Mitarbeit zum Teil als zu aufwändig und zeitintensiv beurteilt wurde. Das wird auch durch das beobachtete Nutzungsverhalten der User/-innen der Plattform unterstrichen. Ein Großteil der Einträge wurde von Wenigen erstellt. Dabei ist besonders zu berücksichtigen, dass die Aktivitäten zur Fachgebärdensammlung in der Freizeit der teilnehmenden Studierenden und Dolmetscher/-innen stattfindet. Daher muss die Verwendung und die Mitarbeit möglichst zeiteffizient und einfach gestaltet werden.

Bei den Rückmeldungen stand besonders eine gute und leichte Möglichkeit zur Erstellung und Einbindung der Videos im Vordergrund. Sie müssen derzeit mit einem externen Gerät aufgenommen werden, auf einem Computer nachbearbeitet und dann in die Plattform eingebunden werden. Dieser Weg erlaubt die Erstellung von Aufzeichnungen unter besseren Rahmenbedingungen. Die Erstellung ist jedoch aufwändiger. Von einigen Benutzern/-innen wurde der Umstand, dass nur ein Video auf einmal hochgeladen werden kann, als zeitintensiv und einschränkend empfunden<sup>12</sup>.

Zur Behebung der vorher angeführten Probleme existieren grundsätzlich zwei Lösungswege. Erstens: mehrere Videos parallel hochzuladen und dabei eventuell auch gleichzeitig mehrere Einträge zu erstellen. Die Realisierung ist schwierig, da die einfache Erstellung multipler Inhal-

<sup>12</sup> Das Hochladen der Videos erfolgt bei der Erstellung oder Bearbeitung eines Eintrages selbst.



(a) Bei dem Video wurde der Hintergrund für die Aufnahme unvoreteilhaft gewählt



(b) Bei dem Video wurde bei der Nachbearbeitung eine zu hohe Komprimierung gewählt

**Abbildung 6.8:** Ausschnitte aus zwei, von Benutzerinnen des GESTU Fachgebärden Wikis, selbsterstellten Videos. Bei der Aufnahme oder Nachbearbeitung wurden nicht alle auf Seite 88 genannten Empfehlungen eingehalten.

te beim Anlagen gelöst werden.

Die zweite und aus Sicht der Benutzer/-innen einfachere Lösung ist es, Videos direkt durch die Plattform aufzunehmen und einzubinden. Dadurch wird sich die Qualität der Aufnahmen wahrscheinlich verschlechtern, wie jedoch auf Seite 89 beschrieben wurde und anhand der Abbildungen 6.8 zu sehen ist, lassen sich dennoch die Gebärden einwandfrei erkennen.

### 6.2.3 Empfehlungen

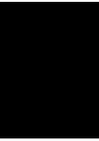
Mit der Erstellung und Umsetzung des GESTU Fachgebärden Wikis wurde ein erster, wichtiger Schritt zur Verringerung der bestehenden Problematik der fehlenden Gebärden für wissenschaftliche Fachausdrücke gemacht. Die geschaffene Plattform stellt dabei eine solide Basis dar. Ihr Erfolg hängt jedoch von der Mitarbeit der Benutzer/-innen und von einer großen Nutzerbasis ab. Bei einer Weiterführung und Weiterentwicklung muss dabei besonders auf diese beiden Faktoren Rücksicht genommen werden.

In den Rückmeldungen der Studierenden zeigten sich einige Schwächen in den gewählten Lösungswegen. Es gibt jedoch bereits Ideen und Vorschläge für konkrete Verbesserungen. An erster Stelle steht dabei die zuvor angesprochene Vereinfachung der Erstellung und Einbindung von Videos direkt unter Verwendung der Plattform. Das würde die Mitarbeit vereinfachen, attraktiver gestalten aber auch die dafür benötigte Zeit verringern. Wird das umgesetzt, verliert die Implementierung der Möglichkeit, mehrere Videos gleichzeitig hochzuladen und mehrere Einträge gleichzeitig zu erstellen, aus meiner Sicht stark an Priorität.

Weiters kann durch die schnellere Bewertung von Gebärdenvorschlägen die Mitarbeit vereinfacht werden. Das kann ähnlich wie beim ASL-STEM Forum durch Vergabe einer Wertung auf einer Skala von eins bis fünf ermöglicht werden (siehe Abschnitt 6.1.4 auf Seite 80).

Die Einbindung des GESTU Fachgebärden Wikis in die Aktivitäten zur Sammlung und Dokumentation von Fachgebärden des Modellversuchs war ein entscheidender Vorteil für die Umsetzung und den Start der Plattform. Die Verwendung in den von Professor Dotter gehaltenen Lehrveranstaltungen mit dem Schwerpunkt der Gebärdensprachlinguistik, war besonders förderlich. Dadurch ergab sich die Chance, direkt und gemeinsam mit einem Teil der Benutzer/-innen eventuelle Änderungen, Verbesserungen und Erweiterungen zu diskutieren. Dem Beispiel folgend sollte das Fachgebärden Wiki auch zukünftig bei eventuellen Kursen mit dem inhaltlichen Schwerpunkt zu Gebärdensprachen verwendet werden. Darüber hinaus könnte die Mitarbeit und Erstellung von Inhalten in die Vorlesung beziehungsweise deren Leistungsbeurteilung integriert werden.

Ein Hindernis zur Schaffung einer größeren Nutzerbasis ist der nicht öffentliche Zugang zu den Einträgen des GESTU Fachgebärden Wikis. Zur Vergrößerung der Nutzeranzahl und Erweiterung der Inhalte sollte eine Öffnung erfolgen. Dazu müssen jedoch die Nutzungsbedingungen angepasst werden. Zusätzlich müssen auch die Autoren/-innen der bestehenden Einträge über die Änderung informiert werden.



# Remote Gebärdensprachdolmetschen

## 7.1 Stand der Technik und Forschung

Videotelefonate in guter Qualität sind heute alltäglich. Sie können nicht nur mit herkömmlichen Mobiltelefonen durchgeführt werden, sondern auch mit Hilfe des Internets und in sehr vielen Fällen kostenlos. Dadurch eröffnen sich besonders für hörbehinderte Personen neue und einfache Möglichkeiten zur Kommunikation, wenn sie räumlich getrennt sind. Neben Gesprächen in einer Gebärdensprache ist es mit Videotelefonaten auch möglich, örtlich entfernte Simultandolmetscher/-innen zur Übersetzung zuzuschalten. Beim sogenannten Remote Gebärdensprachdolmetschen (RGD) werden zu einem Treffen zwischen hörenden und gehörlosen Gesprächspartnern/-innen der/die nicht anwesende Dolmetscher/-in über eine Audio- und Videoverbindung hinzugezogen. Unter Verwendung der Technik kann eine Übersetzung ermöglicht werden, selbst wenn sie sonst nicht möglich wäre. Das ist bei sehr kurzfristig stattfindenden Treffen, bei zu kurzen Terminen oder bei großen Wegzeiten der Fall (siehe [Hat11, S: 84 - 90]). Fotos von Anwendungen des Remote Gebärdensprachdolmetschen sind in Abbildung 7.1 zu sehen.



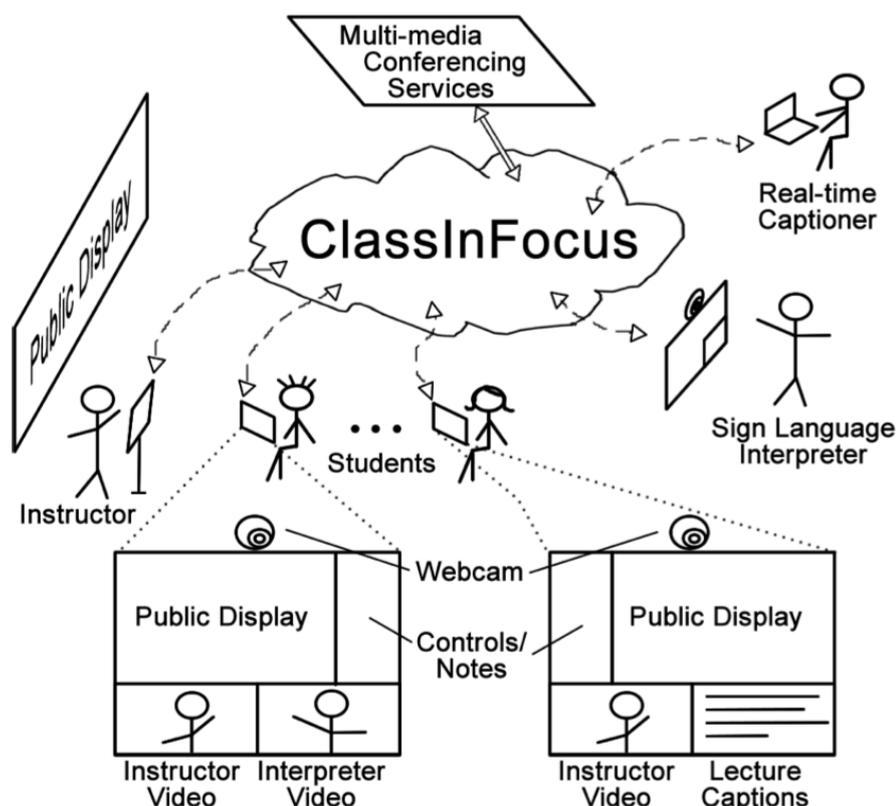
**Abbildung 7.1:** RGD bei zwei möglichen Szenarien [Hat11, Abbildung 3.12].

Bei Videotelefonaten in der Gebärdensprache und im speziellen beim RGD sollte die Verbindung eine gute Qualität aufweisen und bestimmte technische Mindestanforderungen erfüllen. Dabei spielt nicht nur die Verzögerung eine entscheidende Rolle - Hattinger gibt eine maximale Verzögerung von 0,2 Sekunden an (siehe [Hat11, S: 89]) - sondern besonders die Bildqualität. Die Videoauflösung und die Anzahl der Bilder pro Sekunde (FPS steht für Frames per Second) ist in diesem Zusammenhang wesentlich. Hattinger empfiehlt für das RGD zumindest 15 FPS und eine Auflösung von 800 x 600 Pixel, nur in Ausnahmefällen bei schlechter Verbindung 640 x 480 Pixel (siehe [Hat11, S: 154 - 155]). Bei der Kommunikation in einer Gebärdensprache sind kleine Details wie die Mimik aber auch einzelne Fingerstellungen von Bedeutung. Eine schlechte Verbindung hat daher nicht nur Einfluss auf die Verständlichkeit des Inhaltes - im schlimmsten Fall können Gebärden nicht mehr erkannt werden - sondern auch auf die Ermüdung der Teilnehmer/-innen, bedingt durch eine höhere Konzentration zum Verstehen des Inhaltes (siehe [Hat11, S: 84 - 90]). Der zuletzt genannte Punkt betrifft besonders die Gebärdensprachdolmetscher/-innen, selbst wenn ein Videotelefonat in guter Qualität möglich ist:

„Aber selbst wenn gute technische Rahmenbedingungen gegeben sind, kann es dazu kommen, dass bei der Kommunikation über eine räumliche Distanz die (ohne hin sehr komplexe) Tätigkeit des Simultandolmetschens komplexer und schwieriger durchzuführen ist“ [Hat11, S: 87].

Der eingeschränkte Zugriff auf während des Treffens verwendete Materialien, wie zum Beispiel Unterlagen, Skizzen oder Tafelbild, spielt in diesem Zusammenhang eine maßgebliche Rolle. Dadurch wird das Dolmetschen schwieriger und die Qualität der Übersetzung beeinflusst, vor allem wenn während des Vortrages oder im Gespräch darauf Bezug genommen wird. Um dem Problem zu begegnen, ist der Einsatz spezieller Softwareapplikationen zur Anzeige mehrerer Informationsquellen, deren flexiblen Anordnung und Anpassung der Anzeigegröße in Betracht zu ziehen. ClassInFocus ist ein Beispiel für solch ein Programm (siehe [Cav10, S: 29 - 70]). Es wurde speziell für die Verwendung von Echtzeituntertiteln und dem RGD in Lehrveranstaltungen entwickelt. Das Prinzip und der Einsatz von ClassInFocus ist in den Abbildungen 7.2 und 7.3 zu sehen.

Aus den oben angeführten Gründen ist beim RGD auf optimale Arbeitsverhältnisse zu achten. Das schließt nicht nur eine gute Qualität des Videotelefonats, sondern auch den Arbeitsplatz der Dolmetscher/-innen und die Rahmenbedingungen vor und während des Einsatzes - zum Beispiel durch das Bereitstellen sämtlicher verwendeter Unterlagen - mit ein (siehe [Hat11, S: 84 - 90]). Hattinger beschreibt das Remote Gebärdensprachdolmetschens in seiner Diplomarbeit ausführlich (siehe [Hat11, S: 84 - 90, 103 - 115 sowie 124 - 130]).



**Abbildung 7.2:** Eine schematische Darstellung des Prinzips von ClassInFocus [Cav10, Figure 4.1].

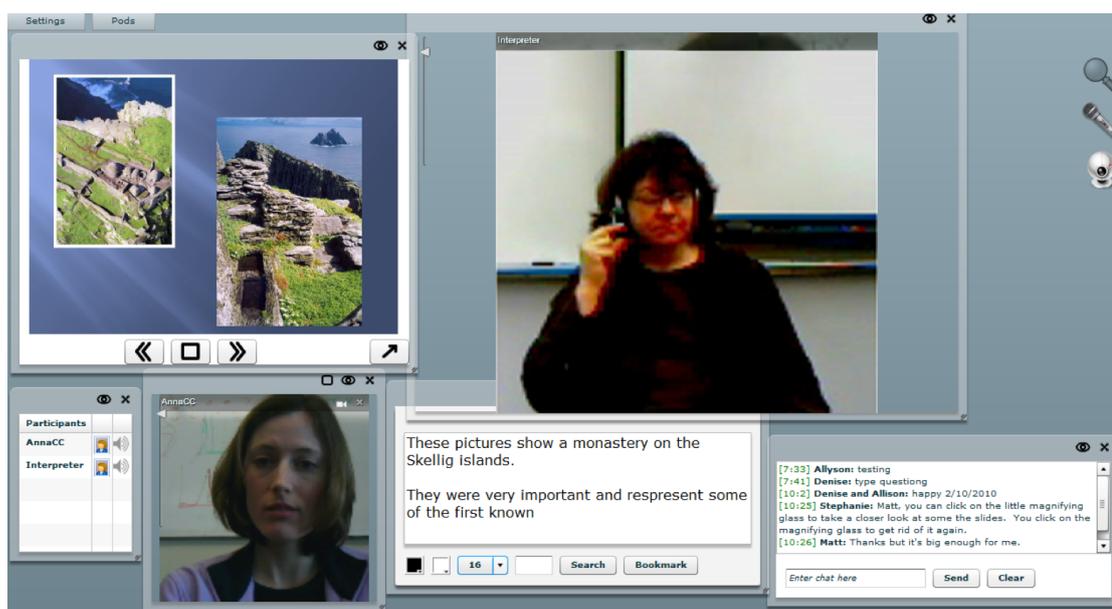


Abbildung 7.3: ClassInFocus im Einsatz [Cav10, Figure 4.2].

## 7.2 Remote Gebärdensprachdolmetschen im GESTU Projekt

Der Einsatz von Gebärdensprachdolmetschern/-innen ist das bevorzugte Mittel, um auditive Informationen für gehörlose Studierende zugänglich zu machen. Im tertiären Bildungssektor bringt das jedoch Probleme mit sich, die zum Teil bereits in den vorhergehenden Kapiteln beschrieben wurden (siehe Kapitel 4.3 auf Seite 32 und Kapitel 6.1.1 auf Seite 74). Die Suche nach qualifizierten Dolmetschern/-innen für sämtliche Lehrveranstaltungen, Nachbesprechungen, Abgabetermine und Prüfungen u.a.m. gestaltet sich oftmals schwierig. Kurzfristige Änderungen oder neu vereinbarte Termine erschweren zusätzlich die Organisation der ÖGS Dolmetscher/-innen. Das Remote Gebärdensprachdolmetschen bietet die Chance, trotz dieser Probleme und für manche Termine dennoch eine Übersetzung in die ÖGS zu ermöglichen. Einer der Hauptaspekte ist dabei der Wegfall der Anfahrtszeiten. Das ermöglicht es Dolmetschern/-innen außerhalb des Großraums Wien, Termine in Wien für das GESTU Projekt wahrzunehmen. Dadurch würden dem Modellversuch mehr Personen für die Übersetzung zur Verfügung stehen. Das RGD ermöglicht weiters durch den Wegfall der Anfahrtszeiten Kosten einzusparen. Zur Erreichung dieses Zieles muss das technische Hilfsmittel jedoch sehr zuverlässig, kostengünstig verfügbar und einfach in der Handhabung sein. Bis zum Start des GESTU Projekts kam das Remote Gebärdensprachdolmetschen im österreichischen tertiären Bildungssektor noch nicht zum Einsatz. Von Hattinger wurde daher eine Fallstudie zur Evaluierung der grundsätzlichen Durchführbarkeit und Brauchbarkeit der Technik durchgeführt. Ausgehend davon empfiehlt er den weiteren Probetrieb in vier aufeinander aufbauenden Phasen. Er rät, das RGD zuerst unter sehr guten und kontrollierbaren Bedingungen zu testen und bei Erfolg nach und nach die Schwierigkeit zu steigern (siehe [Hat11, S: 103 - 115, 124 - 130]). Wenn alle Phasen

erfolgreich durchlaufen werden, soll das Hilfsmittel als Service für gehörlose Studierende im Zuge des GESTU Projekts angeboten und eingesetzt werden.

### 7.2.1 Durchführung

Ausgehend von Hattingers Empfehlungen (siehe [Hat11, S: 124 - 130]) wurde versucht, das Remote Gebärdensprachdolmetschen während des GESTU Projekts weiter zu erproben. Folglich wurde gemeinsam mit den am Modellversuch teilnehmenden Studierenden nach Terminen für weitere Versuche gesucht. Es konnten jedoch nur sehr wenige gefunden werden, die für Versuche mit dem Remote Gebärdensprachdolmetschen geeignet waren. Der Schwerpunkt der Diplomarbeit lag darüber hinaus auf der Erprobung und Evaluierung der in den vorhergehenden Kapiteln beschriebenen technischen Hilfsmittel. Somit kam es im Zuge des GESTU Projekts nur zu einem weiteren Versuch. Es handelte sich dabei um eine ungefähr 1,5 Stunden dauernde Nachbesprechung einer Lehrveranstaltung im Sommersemester 2011. Das Treffen fand in einem Besprechungsraum am Institut für Sprachwissenschaften der Universität Wien statt. Es wurde von einer Professorin für zwei am Modellversuch teilnehmenden gehörlosen Studentinnen abgehalten. Die Dolmetscherin befand sich in der Bibliothek des Instituts „integriert studieren“ an der Technischen Universität Wien.

Bei dem Versuch wurde dieselbe Softwarekonfiguration wie in Hattingers Fallstudie verwendet (siehe [Hat11, S: 103 - 115]). Diesmal wurde jedoch aufgrund seiner Erkenntnisse lediglich Sykpe<sup>1</sup> eingesetzt. Bei der eingesetzten Computerhardware handelte es sich größtenteils um dieselben Geräte, wie sie bei der zuvor genannten Fallstudie verwendet wurden. Im Besprechungsraum wurde ein Laptop - ein 15" MacBook Pro<sup>2</sup> - verwendet, der über das WLAN der Universität Wien mit dem Internet verbunden war. Die Stimme der Professorin wurde mit einem externen Mikrofon<sup>3</sup> aufgezeichnet, wodurch eine bessere Audioqualität erzielt werden konnte. In der Bibliothek wurde der vorhandene Computer<sup>4</sup> in Kombination mit einem 52" Display<sup>5</sup> und einer externen Webcam<sup>6</sup> verwendet, der über das lokale Netzwerk der TU Wien mit dem Internet verbunden war.

### 7.2.2 Erkenntnisse

Der zuvor beschriebene Versuch verlief aufgrund einiger ungünstiger Faktoren nicht optimal. Dadurch sank in der zweiten Hälfte des Treffens die Qualität und Verständlichkeit der Übersetzung. Es konnte dennoch gezeigt werden, dass der Einsatz des RGD grundsätzlich möglich ist, außerdem wurden wesentliche Erkenntnisse betreffend der Rahmenbedingungen gewonnen. Bei den negativen Einflüssen handelte es sich um:

<sup>1</sup> <http://www.skype.com> - letzter Zugriff am 06.03.2013.

<sup>2</sup> Apple Macbook Pro, Intel Core 2 Duo, 2,8 GHz, 4GB Arbeitsspeicher, Betriebssystem: Mac OS X 10.6.4.

<sup>3</sup> Dabei handelte es sich um das bereits für die Videoaufzeichnungen und das Respeaking verwendete Sennheiser ew100 G3 Funkmikrofon Set.

<sup>4</sup> Prozessor Intel Core 2 Duo E8400, 3 GHz, S-775, Boxed, 2 GB Arbeitsspeicher, Betriebssystem: Windows7 Enterprise Edition 32Bit.

<sup>5</sup> Samsung LE-52A656A1F (LE52A656A1FXXC), Auflösung: 1920x1080.

<sup>6</sup> Logitech Webcam C500.

- **Dauer des Versuches**

Ursprünglich wurde die Nachbesprechung mit einer Gesamtdauer von einer Stunde vereinbart. Sie wurde jedoch bei weitem überschritten und nach einer Stunde und 20 Minuten von den beteiligten GESTU Mitarbeitern abgebrochen. Dolmetschereinsätze, die länger als eine Stunde dauern, werden angesichts der hohen Belastungen mit zwei Dolmetschern/-innen, die sich während des Einsatzes abwechseln, besetzt. Wie bereits zu Beginn des Kapitels beschrieben, ist das Simultandolmetschen beim RGD schwieriger und ermüdender als reguläre Übersetzungstermine. Daher hätten für den Versuch zwei Gebärdensprachdolmetscher/-innen eingesetzt werden müssen.

- **Verwendung unterstützender Materialien auf Papier, die der Dolmetscherin nicht zugänglich waren**

Die Professorin verwendete während der Nachbesprechung Notizen und Zeichnungen auf Papier, um Sachverhalte besser darzustellen. Sie waren der Dolmetscherin jedoch nicht zugänglich. Das erschwerte einige Male das Übersetzen während des Versuches. Wie bereits im ersten Kapitel beschrieben, ist es wesentlich, dass solche Materialien dem/der Dolmetscher/-in zugänglich sind, um eine gute Qualität bei der Übersetzung in die Gebärdensprache zu erzielen.

- **Schlechte Videoqualität**

Skype verringerte wegen der zum Teil zu geringen Bandbreite der Internetverbindung oftmals automatisch die Bildqualität des übertragenen Videos. Das geschah trotz der, wie bei Hattinger (siehe [Hat11, S: 107]) beschriebenen, Modifikation der Skype Einstellungen. Die Auflösung des Videos betrug konstant 640 x 480 Pixel. Das ist angesichts der Großaufnahme der Gebärdensprachdolmetscherin noch ausreichend. Während des Videotelefonats wurden maximal zwölf Bilder pro Sekunde übertragen, die öfters auch unter den genannten Wert sanken. Die einzelnen Bewegungen bei der Darstellung der Gebärden waren deshalb verschwommen und schwer, beziehungsweise zum Teil gar nicht erkennbar.

Für die Verständlichkeit der Übersetzung war infolgedessen ein erhöhtes Maß an Konzentration erforderlich. Bei dem Versuch zeigte sich weiters, dass für die gute Verständlichkeit der Übersetzung in die Gebärdensprache eine Untergrenze für die Bilder pro Sekunde (FPS) wesentlich ist. Sie sollte am besten manuell auf mindestens 20 FPS festgelegt werden können.

Neben den negativen Auswirkungen auf die beiden gehörlosen Studentinnen war der Versuch für die Dolmetscherin sehr fordernd und anstrengend. Das zeigte sich besonders in der zweiten Hälfte der Nachbesprechung. Das Ergebnis ist dennoch nicht als Fehlschlag beziehungsweise als Zeichen für eine Unbrauchbarkeit des RGD zu werten. Vielmehr zeigte sich, dass beim Einsatz vorerst noch umsichtig vorgegangen werden muss. Es empfiehlt sich, zukünftig in Zusammenarbeit mit allen Beteiligten, besonders aber den Dolmetschern/-innen, möglichst optimale Bedingungen sicherzustellen. Dabei soll besonders der Zeck, der voraussichtliche Inhalt und die Dauer des geplanten Termins festgestellt und verwendete Materialien ausgetauscht werden. Für den alltäglichen Einsatz des RGD sind noch weitere Versuche durchzuführen. Sie sollten

zunächst, wie auch von Hattinger vorgeschlagen, unter einfacheren Rahmenbedingungen stattfinden (siehe [Hat11, S: 124 - 127]).

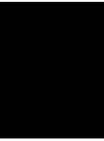
Hinsichtlich der Einsatzmöglichkeiten zeigte sich während des GESTU Projekts ein neues Anwendungsszenario. In der zweiten Hälfte des Modellversuchs wurden mehrere Dolmetscherinnen fest angestellt und nicht nur für einzelne Termine und Lehrveranstaltungen gebucht auf Honorarbasis. Sind sie zu bestimmten Zeiten in der Servicestelle anwesend, ist es denkbar, das RGD zu festgelegten Zeiten als flexibles und kurzfristig nutzbares Service anzubieten. Wird das angeboten, ist es sinnvoll, einen speziellen Bereich für das Remote Gebärdensprachdolmetschen einzurichten. Somit können gute Arbeitsbedingungen für die Dolmetscher/-innen bei der Verwendung der Technik geschaffen werden. Hattinger beschreibt in seiner Diplomarbeit die Gestaltung einer solchen Arbeitsumgebung (siehe [Hat11, S: 128 - 129]).

### 7.2.3 Empfehlungen

Das Remote Gebärdensprachdolmetschen ist grundsätzlich im österreichischen tertiären Bildungssektor einsetzbar, wie durch die von Hattinger (siehe [Hat11, S: 103 - 115]) durchgeführte Fallstudie aber auch den zweiten, auf Seite 97 beschriebenen Versuch, gezeigt wurde. Gelingt es, das technische Hilfsmittel einfach und ohne ständige Betreuung durch technische Mitarbeiter/-innen einzusetzen, bietet es zahlreiche neue Möglichkeiten zum flexibleren Einsatz von Simultandolmetschern/-innen. Um das zu erreichen, sind jedoch noch weitere Versuche, zunächst unter besseren Rahmenbedingungen, notwendig. Für die weitere Vorgangsweise bieten sich die Empfehlungen von Hattinger an (siehe [Hat11, S: 124 - 127]). Bei der weiteren Erprobung ist es wesentlich, die Problematik der unzureichenden Bildqualität zu verbessern. In dem Zusammenhang spielt die Findung und der Einsatz einer Videokonferenz- oder Videotelefoniesoftware, bei der die einzelnen Parameter des Videos - im Speziellen die Bildrate - einstellbar sind, eine große Rolle. Das Problem der für die Dolmetscher/-innen nicht zugänglichen Informationsquellen, die während des Remote Gebärdensprachdolmetschens verwendet werden, kann eventuell durch den Einsatz spezieller Softwareapplikationen - wie zum Beispiel ClassInFocus (siehe [Cav10, S: 29 - 70]) - entschärft werden. Das sollte auch bei weiteren Versuchen mit dem RGD berücksichtigt werden.



KAPITEL 8



**Schlusswort und Ausblick**

In der Diplomarbeit wurden verschiedene technische Hilfsmittel ausgewählt und im GESTU Projekt erprobt. Sie wurden in Zusammenarbeit mit den hörbehinderten Teilnehmern/-innen evaluiert. Die einzelnen Maßnahmen wurden zum Teil sehr unterschiedlich wahrgenommen und bewertet. Das lag an den einzelnen Hilfsmitteln, den Rahmenbedingungen für den Einsatz und an den unterschiedlichen Bedürfnissen der einzelnen Studierenden. Das zeigte sich besonders bei der Echtzeit-Untertitelung mittels der Respeaking Technik, aber auch bei der Lehrveranstaltungsaufzeichnung und dem GESTU Fachgebärden Wiki. Bereits Hattinger hebt die oben angeführten Aspekte beim Einsatz technischer Hilfsmittel für hörbehinderte Studenten/-innen im Schlusswort seiner Diplomarbeit hervor:

„Bei Studierenden mit Hörbeeinträchtigung handelt es sich um eine sehr heterogene Gruppe. Da auch universitäre Veranstaltungen sehr vielfältig sind - abhängig vom Studienfach, der Art der Lehrveranstaltung, den beteiligten Personen - müssen verschiedene Hilfsmittel den jeweiligen Bedingungen gerecht werden“ [Hat11, S: 141].

Für die einzelnen technischen Hilfsmittel wurden im jeweiligen Kapitel bereits die Erkenntnisse und eine Empfehlung für den zukünftigen Einsatz und die Weiterentwicklung beschrieben. Abschließend werden die Ergebnisse nachfolgend nochmals zusammengefasst.

Die **Lehrveranstaltungsaufzeichnung für gehörlose Studierende** wurde am meisten eingesetzt. Sie wurde von den Teilnehmern/-innen des GESTU Projekts sehr gut bewertet und wahrgenommen. Bei den Befragungen zeigte sich, dass die Aufzeichnungen eine große Hilfe beim Erlernen und Vertiefen des Stoffes der Vorlesungen darstellen. Zusätzlich erlaubten sie es sprachkulturelle Barrieren, die für gehörlose Studierende an lautsprachlichen Bildungseinrichtungen bestehen, zu verringern. Sofern die Aufnahmen auch den anderen Besuchern/-innen der Kurse zur Verfügung gestellt werden, können auch sie davon profitieren.

Mit einem Teil der Lehrveranstaltungsaufzeichnungen wurde auch versucht, ein **sprecherunabhängiges Spracherkennungssystem für freie Sprache für den österreichischen Dialekt zu trainieren**. Damit könnten automatisiert Untertitel für beliebige Aufnahmen von Vorlesungen des österreichischen tertiären Bildungssektors erstellt werden. Mit dem durchgeführten Training konnten jedoch noch keine ausreichenden Ergebnisse erzielt werden. Für bessere Erkennungsraten ist ein wesentlich größerer Aufwand notwendig.

Neben den Versuchen mit der nachträglichen Untertitelerzeugung wurden auch Untertitel, die in Echtzeit erstellt wurden, in Vorlesungen erprobt. Sie wurden durch die **Respeaking** Technik erstellt, die bis jetzt noch nicht in Lehrveranstaltungen des österreichischen tertiären Bildungssektor verwendet wurde. Neben der grundsätzlichen Einsetzbarkeit der Technik im Alltagsbetrieb an den Wiener Universitäten zeigte sich, dass das Respeaking nur für bestimmte Vorlesungstypen und Inhalte geeignet ist. Werden gute Rahmenbedingungen geschaffen, stellen die Live-Untertitel für manche Studierende eine sehr gute Alternative zur Übersetzung in die Gebärdensprache dar. Dadurch kann die Problematik der zu geringen Anzahl an ÖGS Dolmetschern/-innen in Wien entschärft werden.

Ein weiterer Schwerpunkt der Diplomarbeit war die Konzeptionierung und Betreuung einer **Internet-Plattform zur Sammlung von Fachgebärden**. Das sogenannten **GESTU Fachgebärden Wiki** konnte schnell und kostengünstig realisiert werden. Nach der Fertigstellung wurde

es in der zweiten Hälfte des Projekts in diverse Aktivitäten des Modellversuchs eingebunden. Dadurch konnte in kurzer Zeit ein vergleichsweise großer Grundstock an Einträgen gesammelt werden. In den Rückmeldungen der Benutzer/-innen zeigten sich einige Schwächen in den gewählten Lösungswegen. Denkbare Verbesserungen sind eine einfachere Möglichkeit zur Mitarbeit und zur Aufnahme der Videos. Neben deren Umsetzung stellt auch die Vergrößerung des Benutzerkreises den nächsten Schritt dar.

Im Zuge der Diplomarbeit wurden weitere Versuche mit dem **Remote Gebärdensprachdolmetschen** angestrebt. Aufgrund mangelnder geeigneter Termine kam es jedoch nur zu einem weiteren Experiment. Mehrere negative Einflussfaktoren führten - besonders in der zweiten Hälfte des Versuchs - zu schlechten Ergebnissen. Das ist jedoch nicht als generelles Versagen der Technik zu werten. Es sind weitere Versuche und das Sicherstellen guter Rahmenbedingungen bei der Verwendung notwendig, um das Remote Gebärdensprachdolmetschen im alltäglichen Universitätsbetrieb zuverlässig einsetzen zu können.

Die Erprobung und der Einsatz der beschriebenen technischen Hilfsmittel im Zuge der Diplomarbeit während des GESTU Projekts gewährte neue Erkenntnisse zur Unterstützung hörbehinderter Studierender. Viele der ausgewählten Maßnahmen konnten nahtlos in den täglichen Universitätsbetrieb integriert werden und stellten für die Teilnehmer/-innen des Modellversuchs einen Beitrag zum Abbau bestehender Hürden dar.



KAPITEL 9

**Anhang**

## 9.1 Transkriptvergleiche einer ausgewählten Vorlesung

Im vorliegenden Abschnitt des Anhangs werden die Transkripte der ersten fünf Minuten der Vorlesung „Trainingswissenschaft“ vom 05.06.2012, die durch verschiedene Erzeugungsmethoden produziert wurden, dem manuell erstellten 1:1 Text des Vortrages gegenübergestellt. Zuerst erfolgt der Vergleich mit dem Ergebnis der trainierten EML ASR (siehe Kapitel 3.2). Dabei zeigen sich deutlich die Auswirkungen der schlechten Erkennungsraten.

Beim zweiten Text handelt es sich um Live-Untertitel der Vorlesung, die von der Firma Titelbild durch die Respeaking Technik erstellt wurden (siehe Kapitel 4.3). Bei der Übersetzung in der Lehrveranstaltung kam es vermehrt zu Auslassungen und Zusammenfassungen. Sie waren größtenteils durch das mitunter schnelle Sprechtempo des Vortragenden und die lückenhaft bereitgestellten Unterlagen bedingt. Hattinger hat für Ausschnitte des Vorlesungstermins stichprobenartig die NER Werte berechnet. Dabei wurde im Durchschnitt einen Wert von ~95% erreicht. Das Ergebnis liegt zwar unter den von Romero-Fresco geforderten 98% ist aber dennoch ein gutes Resultat (siehe [Hat13, Kapitel 4], [RFM13]).

### 9.1.1 Das von der trainierten EML ASR erstellte Transkript

| Transkript der EML ASR  | 1:1 Transkript (manuell erstellt)   |
|---|---|
| <p>und dann gibt es das nächste Kapitel des Gebietes Wettkämpfe dass es beschäftigen wird sogar gewertet in der letzten Einheit in 14 Tagen</p>   | <p>Und dann gibt es noch das nächste Kapitel. Das Kapitel Wettkämpfe. Das uns heute beschäftigen wird. Und möglicherweise dann in der letzten Einheit in 14 Tagen.</p>  |
| <p>und zum Schluss wird derzeit noch offen bleibt überlebt kommen Bemerkungen zur Thematik aus der Konzepte nur und besonders belegte des Kindertrainings</p>   | <p>Und zum Abschluss, je nachdem wie viel Zeit noch offen bleibt, übrig bleibt, kommen noch ein paar Bemerkungen zur Thematik Ausdauerkonzepte-neu und besondere Aspekte eines Klimatrainings.</p>  |
| <p>für für aber hat die Übernahme daran glaubt verschiedener Sportarten Hinblick auf das eigene diesen letzten gegen werden oder auch wieder mit verspielen geht dann in entsprechender Weise stattfinden oder und ein Sprecher mitwirken hatte Bedingungen beginnt und natürlich verbittert sind so wird das wir das gesamte begrenzte Wettkampfvorbereitung den Wettkampf oder sich genauer anschauen</p> | <p>Das sind also Dinge, die man wahrscheinlich nicht immer hört. Die aber momentan gerade in verschiedenen Sportarten in Hinblick auf Großereignisse, wie sie in den letzten Jahren mit Peking waren oder auch immer wieder mit Winterspielen, die dann in entsprechender Höhenlage stattfinden oder unter entsprechenden mikroklimatischen Bedingungen, wie in Peking, natürlich von Bedeutung sind. Insoweit, dass man das gesamte Paket der Wettkampfvorbereitung, den Wettkampfort betreffend, sich genauer anschaut.</p> |

das Kürzel zu aber das ist immer mehr der werden über das Thema der Stelle und Trainingsplanung oder zur wenigstens durch Wiederherstellung der Trainingsplanung gewissermaßen ein Paket zu konservativ oder erschlagen wird diese Darstellung von Professor war

das natürlich der Trainingsplanung der komplexen der ist man allerdings ging zur Feinheiten zu unterscheiden und hat das Ganze insoweit zu Kapitel aufgeteilt werden aber das technisch Trainingswissenschaft der SPD für den Trainer war besonderer Abschnitte zu beurteilen ist zu und die muss man natürlich immer in der lettischen Konzept bei und Sprecher der wird dann kann man wird es bleibt ist das ist dann wieder zusammen fließen lassen

oder sogar der Planung wenn sich der Staat von Beginnern normale jetzt mit einer Bestimmung des Ist-Zustand wird ist der Ist-Zustand für die Planung und findet nicht inhaltlich 10.

da haben 2 Kernbereiche die uns zu Beginn seiner erarbeiten müssen diese sondern die Fragestellungen der Welt ist ein Athlet gesund

ist die Stadt für die Basisuntersuchungen wird der Planung zumindest bis Sportarten die so ein Sprecher konditionsintensiven Sehenswertes ganze 11. bis werde es dreimal

dass das oder wichtige Basisuntersuchungen die vor das Geld das ist betreffen und verschiedener Parameter geht den Sonderstatus Planung bestreiten oder bestätigen und das ist jetzt hier und da Belastbarkeits diagnostik angeführt der Bereiche

Ja und ganz kurz zur Erinnerung an das letzte Mal. Wir haben uns beim letzten Mal unterhalten über das Thema Trainingssteuerung und Trainingsplanung. Ich hab dazu erwähnt, dass eigentlich die Trainingssteuerung und die Trainingsplanung gewissermaßen ein Paket sind und wenn wir auf die Folie drauf schauen, auf diese Darstellung von Professor Baron.

Dann sehen wir natürlich dass die Trainingsplanung ein Teil der komplexen Trainingssteuerung ist - „nonanet“. Allerdings gibt es hier einige Feinheiten zu unterscheiden und ich habe das Ganze insoweit in zwei Kapitel aufgeteilt, weil rein arbeitstechnisch für den Trainingswissenschaftler respektive für den Trainer besondere Abschnitte zu beurteilen sind und die muss man natürlich einmal im theoretischen Konzept verinnerlichen und entsprechend aufarbeiten und dann kann man in der Trainingspraxis das System dann wieder zusammenfließen lassen.

Gut sowohl die Planung wie die Trainingssteuerung beginnen einmal - „nonanet“ - mit einer Bestimmung des Istzustandes, wobei klarerweise der Istzustand für die Planung und für die Trainingssteuerung inhaltlich ident sind.

Da haben wir zwei Kernbereiche, die wir uns zu Beginn erarbeiten müssen. Das ist einmal die Fragestellung, inwieweit ist mein Athlet gesund.

Das ist die Sportärztliche-Basisuntersuchung, die im Rhythmus der Planung zumindest jährlich passiert. In Sportarten, die also entsprechend konditionsintensiv sind, wird das ganze öfter passieren - 2-3x.

Das sind also immer wieder wichtige Basisuntersuchungen, die vor allem das Herzkreislaufsystem betreffen und verschiedene Parameter die den Gesundheitsstatus dann unterschreiten oder eben bestätigen und was immer wieder vergessen wird, hier unter Belastbarkeitsdiagnostik angeführt. Das ist der Bereich hier.

besichtigen Fragestellungen da ist man letztendlich betreibe aber wirklich gesund und geht es um den orthopädischen Gesundheitszustand und also wieder auch genannt die Auswertung von Großereignissen ist dann kann man hört damals wirklich versteht dass viele Athleten ist so vor einer Personal für die Koalition zusammen und dann von der ist der Qualifikation aufgrund von beinahe Verletzungen und dann spricht mich Ausfall

umgekehrt kann es passieren dass der Sportler Krippenplätze Überzeugungen haben und dann ist man Deutschland müssen wir gesund nicht in der Lage sind derart große Turnier dann zu bestehen

von der Situation geringer ist man immer ganz sicher bis jetzt die Spiele der Gegend werden wo früher aus ich werde sollten und das ist doch der die Spiele ist wurde das stecken

ist also ganz böse Schlagzeilen Zeitungen über unseren Spielerin offensichtlich an haben wir uns dann Problematik gelitten hatte offensichtlich wird

und die Situation und das schon gehört wurde ist das Bestehen eines wird das 1. passen würde wenn stellen Formate Gesundheitszustand nicht möglich ist und er ist der 1. Van wird nicht dass wir für wirklich stellen und darf wurde die gesperrt sozusagen spricht ausgeschlossen

wird das natürlich die Betroffene spart absolut Peter Situation und umgekehrt belegt so dass das offensichtlich gehört und dann stecken kann Sportarten dieses Zusammenwirken Mannschaft oder als Trainer Trainings Umfeld das Trainer-Berater-System mit einer ausreichend qualitativ hochwertige und läuft

Das ist die Grundfragestellung: inwieweit ist mein Athlet, den ich betreue auch wirklich gesund. Und da geht es um den orthopädischen Gesundheitszustand und wenn man also immer wieder im Nachhinein die Auswertungen von Großereignissen liest, dann kann man dort sehr häufig feststellen, dass viele Athleten, die zu Großereignissen fahren, ich habe es das letzte Mal angesprochen, eigentlich gar nicht gesund sind und dann eigentlich im Vorfeld bereits bei der Qualifikation aufgrund von Banalverletzungen unter Anführungsstrichen scheinbar ausfallen.

Umgekehrt kann es passieren, dass Sportler Quotenplätze bereits errungen haben und dann im letzten Moment ausscheiden müssen, weil sie eben gesundheitlich nicht in der Lage sind, ein derart großes Turnier dann zu bestehen.

Und wer sich da zurück erinnert, ich bin mir nicht mehr ganz sicher ob das jetzt die Spiele in Peking waren, wo im Vorfeld aus einem Beachvolleyball Team durch den Olympiarzt, durch den Doktor Engel, eine Spielerin sozusagen eliminiert wurde, unter Anführungsstrichen.

Das hat also ganz böse Schlagzeilen in den Zeitungen gebracht. Wo also eine Spielerin offensichtlich an einem Wirbelsäulenproblematik gelitten hat, offensichtlich noch immer leidet.

Und die Situation vom Olympiarzt eben so beurteilt wurde, dass das Bestehen eines Olympiaturniers inklusive aller Strapazen die dort entstehen. Vom aktuellen Gesundheitszustand her nicht möglich ist und er als Olympiarzt die Verantwortung nicht übernimmt, dass hier Folgeschäden entstehen und daraufhin wurde die Sportler sozusagen Anführungsstrichen ausgeschlossen.

Gut das ist natürlich für die betroffene Sportlerin eine absolut bittere Situation. Umgekehrt belegt also das genau das, dass offensichtlich gerade in unter Anführungsstrichen Randsportarten dieses Zusammenwirken Mannschaftsarzt, Trainer, Trainingsumfeld, also Trainer-Beratersystem nicht immer ausreichen qualitativ hochwertig abläuft.

### 9.1.2 Das mithilfe der Respeaking Technik erstellte Transkript

| Respeaking Transkript  | 1:1 Transkript (manuell erstellt)  |
|--|--|
| <p>Dann kommen wir zum nächsten Kapitel: Wettkämpfe. Das beschäftigt uns heute und möglicherweise in der letzten Einheit in 14 Tagen.</p> <p>Zum Abschluss, je nachdem wie viel Zeit sie noch haben, noch ein paar Bemerkungen zu besonderen Aspekten eines klimatrainingsniveau. Dabei geht es um Training unter besonderen klimatischen Bedingungen.</p> | <p>Und dann gibt es noch das nächste Kapitel. Das Kapitel Wettkämpfe. Das uns heute beschäftigen wird. Und möglicherweise dann in der letzten Einheit in 14 Tagen.</p> <p>Und zum Abschluss, je nachdem wie viel Zeit noch offen bleibt, übrig bleibt, kommen noch ein paar Bemerkungen zur Thematik Ausdauerkonzepte-neu und besondere Aspekte eines Klimatrainings. Das sind also Dinge, die man wahrscheinlich nicht immer hört. Die aber momentan gerade in verschiedenen Sportarten in Hinblick auf Großereignisse, wie sie in den letzten Jahren mit Peking waren oder auch immer wieder mit Winterspielen, die dann in entsprechender Höhenlage stattfinden oder unter entsprechenden mikroklimatischen Bedingungen, wie in Peking, natürlich von Bedeutung sind. Insoweit, dass man das gesamte Paket der Wettkampfvorbereitung, den Wettkampfort betreffend, sich genauer anschaut.</p> |
| <p>Trainingssteuerung und die Trainingsplanung. Zur Erinnerung: das letzte Mal sprachen wir über die Das ist gewissermaßen ein Paket.</p>  | <p>Ja und ganz kurz zur Erinnerung an das letzte Mal. Wir haben uns beim letzten Mal unterhalten über das Thema Trainingssteuerung und Trainingsplanung. Ich hab dazu erwähnt, dass eigentlich die Trainingssteuerung und die Trainingsplanung gewissermaßen ein Paket sind</p>  |
| <p>Auf der Folie sieht man, dass das die Trainingsplanung ein Teil der komplexen Trainingssteuerung ist. Es gibt aber Feinheiten zu unterscheiden.</p>   | <p>und wenn wir auf die Folie drauf schauen, auf diese Darstellung von Professor Baron. Dann sehen wir natürlich dass die Trainingsplanung ein Teil der komplexen Trainingssteuerung ist - „nonanet“. Allerdings gibt es hier einige Feinheiten zu unterscheiden</p>   |

Ich habe es in zwei Kapitel aufgeteilt, weil ein arbeitstechnisch besonderer Abschnitte zu beurteilen sind. Die müssen im theoretischen Konzept der inner nicht werden. Dann kann man in der Praxis das wieder zusammen fließen lassen.

Sowohl die Planung wie die Steuerung beginnen mit einer Bestimmung des Istzustands.

Wir haben hier zwei Kernbereiche, die wir uns ansehen müssen. Zum einen die Fragestellung: inwieweit ist mein Athlet gesund? Das muss mindestens jährlich passieren. Bei konditionsintensiven Trainings geschieht das zwei bis dreimal im Jahr.

Das sind wichtige Basisuntersuchungen des Herzkreislaufsystems und anderen Parametern des Gesundheitsstatus. Die Grundfragestellung, inwieweit der Athlet wirklich gesund ist, wird immer wieder vergessen. Hier geht es um den orthopädischen Gesundheitszustand.

Bei der Auswertung von Großereignissen, sieht man sehr oft, dass viele Athleten, die zu diesen Ereignissen fahren, gar nicht gesund sind. Sie fallen dann scheinbar schon im Vorfeld aus, aufgrund von banalen Verletzungen.

und ich habe das Ganze insoweit in zwei Kapitel aufgeteilt, weil rein arbeitstechnisch für den Trainingswissenschaftler respektive für den Trainer besondere Abschnitte zu beurteilen sind und die muss man natürlich einmal im theoretischen Konzept verinnerlichen und entsprechend aufarbeiten und dann kann man in der Trainingspraxis das System dann wieder zusammenfließen lassen.

Gut sowohl die Planung wie die Trainingssteuerung beginnen einmal - „nonanet“ - mit einer Bestimmung des Istzustandes, wobei klarerweise der Istzustand für die Planung und für die Trainingssteuerung inhaltlich ident sind.

Da haben wir zwei Kernbereiche, die wir uns zu Beginn erarbeiten müssen. Das ist einmal die Fragestellung, inwieweit ist mein Athlet gesund. Das ist die Sportärztliche-Basisuntersuchung, die im Rhythmus der Planung zumindest jährlich passiert. In Sportarten, die also entsprechend konditionsintensiv sind, wird das ganze öfter passieren - 2-3x.

Das sind also immer wieder wichtige Basisuntersuchungen, die vor allem das Herzkreislaufsystem betreffen und verschiedene Parameter die den Gesundheitsstatus dann unterschreiten oder eben bestätigen und was immer wieder vergessen wird, hier unter Belastbarkeitsdiagnostik angeführt. Das ist der Bereich hier. Das ist die Grundfragestellung: inwieweit ist mein Athlet, den ich betreue auch wirklich gesund. Und da geht es um den orthopädischen Gesundheitszustand

und wenn man also immer wieder im Nachhinein die Auswertungen von Großereignissen liest, dann kann man dort sehr häufig feststellen, dass viele Athleten, die zu Großereignissen fahren, ich habe es das letzte Mal angesprochen, eigentlich gar nicht gesund sind und dann eigentlich im Vorfeld bereits bei der Qualifikation aufgrund von Banalverletzungen unter Anführungsstrichen scheinbar ausfallen.

Manche Sportler können auch das letzte große Turnier nicht bestehen, weil sie Verletzungen haben. Bei den Spielen in Peking wurde eine Spielerin eliminiert, die eine Krankheit hatte. Der Arzt sagte, sie kann diesen Wettkampf nicht durchführen. Daraufhin wurde sie ausgeschlossen. Das war für die Sportlerin sehr bitter.

Es belegte aber auch, dass das Zusammenwirken zwischen Training, Trainer, Umfeld nicht immer qualitativ hochwertige ausfällt.

Umgekehrt kann es passieren, dass Sportler Quotenplätze bereits errungen haben und dann im letzten Moment ausscheiden müssen, weil sie eben gesundheitlich nicht in der Lage sind, ein derart großes Turnier dann zu bestehen. Und wer sich da zurück erinnert, ich bin mir nicht mehr ganz sicher ob das jetzt die Spiele in Peking waren, wo im Vorfeld aus einem Beachvolleyball Team durch den Olympiaarzt, durch den Doktor Engel, eine Spielerin sozusagen eliminiert wurde, unter Anführungsstrichen. Das hat also ganz böse Schlagzeilen in den Zeitungen gebracht. Wo also eine Spielerin offensichtlich an einem Wirbelsäulenproblematik gelitten hat, offensichtlich noch immer leidet. Und die Situation vom Olympiaarzt eben so beurteilt wurde, dass das Bestehen eines Olympiaturniers inklusive aller Strapazen die dort entstehen. Vom aktuellen Gesundheitszustand her nicht möglich ist und er als Olympiaarzt die Verantwortung nicht übernimmt, dass hier Folgeschäden entstehen und daraufhin wurde die Sportler sozusagen Anführungsstrichen ausgeschlossen. Gut das ist natürlich für die betroffene Sportlerin eine absolut bittere Situation.

Umgekehrt belegt also das genau das, dass offensichtlich gerade in unter Anführungsstrichen Randsportarten dieses Zusammenwirken Mannschaftsarzt, Trainer, Trainingsumfeld, also Trainer-Beratersystem nicht immer ausreichen qualitativ hochwertig abläuft.

## **9.2 Leitfäden für die eTutoren/-innen**

Nachfolgend sind die drei Leitfäden zur Durchführung des Respeakings und der Lehrveranstaltungsaufzeichnung im Zuge des GESTU Projekts enthalten. Sie wurden den eTutoren/-innen bei ihren Tätigkeiten als Hilfestellung zur Verfügung gestellt. Darin sind alle erforderlichen Schritte für die einzelnen Maßnahmen beschrieben.

Bei den hier abgedruckten Leitfäden wurden alle Zugangsdaten und Namen entfernt.

### 9.2.1 eTutor Leitfaden für die Durchführung von Respeaking



## Respeaking

Dieser Leitfaden dient zur Übersicht und als Hilfestellung für die Vorgehensweise vor und während der Live-Untertitelung durch die Respeaking Technologie.

Bei Fragen Werner Nemecek unter 01/234 567 kontaktieren.

### Vorbereitung

- Folien und Skripte an die Firma Titelbild senden.
- Die Übersendung sollte spätestens drei Tage vor dem Vorlesungstermin erfolgen.
- Equipment überprüfen und vorbereiten.
- Rechtzeitig vor dem Vorlesungsbeginn eintreffen, um einen geeigneten Platz wählen zu können (in der Regel sind 15 Minuten ausreichend).

### Vor Vorlesungsbeginn

#### Mikrofon-Set überprüfen und einstellen

- Verfügen die Akkumulatoren in beiden Geräten über eine ausreichende Aufladung? Sender wie Empfänger müssen mindestens zwei Striche anzeigen.
- Ist die „Sensitivity“ beim Sender auf -18dB eingestellt?
- Ist „AF Out“ beim Empfänger auf 0dB eingestellt?
- Empfang beim Empfänger - auch während der Aufnahme - überprüfen.

Den Ausgang der Empfangseinheit (3,5mm Klinke Stecker) beim Mikrofoneingang des Laptops anstecken.

### Software

#### Respeakingssession im Browser

Zum Login die folgenden Parameter auf der Seite der Firma Titelbild eingeben:

- URL: <http://AdresseZumRespeakingsservice.com>
- eMail-Address: Zugang
- Password: dein Passwort

Wenn der Login erfolgreich war, erscheint eine Übersicht mit den abgelaufenen (grauen) Sessions sowie der aktiven (grünen) Session. Der aktiven Session beitreten (anklicken). Die aktive Session wird in einem neuen Browser-Fenster geöffnet.

**Skype**

User: dein Benutzername

Passwort: dein Passwort

- Überprüfen ob in den Skype-Einstellungen der richtige Audio-Eingang ausgewählt ist.
- Dem/der Vortragenden das Mikrofon befestigen - Brust Mitte, an nicht allzu beweglichen Kleidungsstücken (z.B.: am Sakko, der Hemdtasche,...) – und die Sendeeinheit übergeben (+Mute ausschalten).
- Auf Aufzeichnungsbedingungen hinweisen - Meldungen der Studierenden sollen vom/von der Vortragenden wiederholt werden. Eventuelle Fragen des/der Vortragenden klären.
- Die Audioverbindung mit der Firma Titelbild herstellen.
- Mit der Firma Titelbild ebenfalls nochmals einen Soundcheck durchführen. Hierbei darauf achten, dass der Eingang nicht zu laut oder zu leise ist oder ob Skype die Aufnahmelautstärke regelt.

**Während der Vorlesung**

- Überprüfung der ordnungsgemäßen Funktion von Laptop und Mikrofon-Empfangseinheit achten.
- Der/die Studierende kann folgende Einstellungen in der Webapplikation von Titelbild mit dem ausklappbarem Menü links ändern:
  - Zeilenumbruch
  - Text-Scrollgrenze
  - Textgröße
  - Satznummerierung

**Nach der Vorlesung**

- Equipment wieder abbauen.
- Untertitel aus dem Browser speichern und dem/der Studierenden via E-Mail übersenden.
- Übermittlung des Transkriptes an die/den Vortragende/-n, sofern es gewünscht wird.

### 9.2.2 eTutor Leitfaden für die Lehrveranstaltungsaufzeichnung



#### Lehrveranstaltungsaufnahme

Der Leitfaden dient zur Übersicht und als Hilfestellung für die Vorgehensweise vor und während der Lehrveranstaltungsaufnahme. Der Ablauf und die benötigten Geräte variieren bei den einzelnen Lehrveranstaltungen und Universitäten geringfügig.

Bei Fragen Werner Nemecek unter 01/234 567 kontaktieren.

#### Inhalt

|   |          |
|---|----------|
| <b>1. Vorbereitung .....</b>  | <b>2</b> |
| <b>2. Manuelle Aufzeichnung .....</b>   | <b>2</b> |
| <b>3. Aufzeichnung mit einer im Hörsaal installierten LectureTube - Aufzeichnungseinheit der TU Wien.....</b> | <b>5</b> |

## 1. Vorbereitung

- Ausreichende Aufladung des Akkus der Kamera?
- Ist das Speichermedium der Kamera formatiert oder befinden sich darauf noch alte (eventuell ungesicherte Aufnahmen)?
- Wurden die Reservebatterien für das Funk-Mikrofonset aufgeladen und befinden sie sich im Rucksack?
- Rechtzeitig vor Vorlesungsbeginn eintreffen um einen geeigneten Platz wählen zu können (in der Regel sind 15 Minuten ausreichend).

## 2. Manuelle Aufzeichnung

### Vor Vorlesungsbeginn

#### Suche eines geeigneten Platzes

Die folgenden Punkte sollten bei der Platzwahl berücksichtigt werden:

- Mindestabstand vom Filmobjekt: 2m.
- Wenn Dolmetscher/-innen in der Vorlesung anwesend sind, einen Platz in der 1. - 3. Reihe wählen. Die Kamera sollte idealerweise auf Augenhöhe platziert werden können.
- Bei Aufnahmen ohne Dolmetscher/-innen Wahl eines Platzes, sodass genug Abstand zu den Folien und dem/der Vortragenden besteht.
- Bei Hörsälen ohne oder mit flach steigenden Reihen am besten einen Platz in der erste Reihe wählen.

### Kamera und Stativ vorbereiten

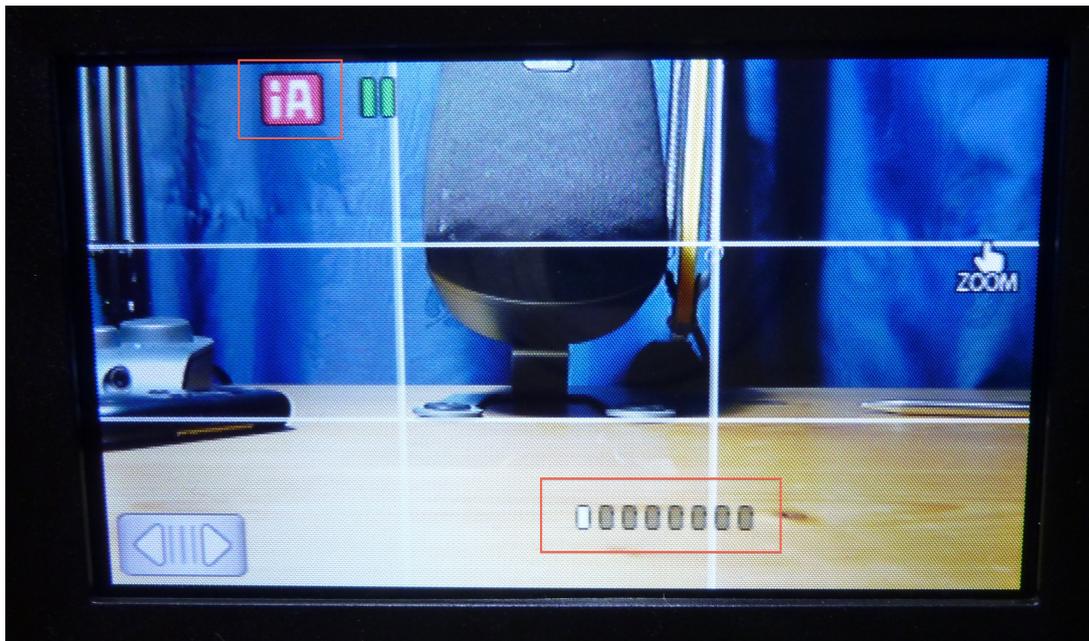
#### Mikrofon-Set überprüfen und einstellen

- Verfügen die Akkumulatoren in beiden Geräten über eine ausreichende Aufladung. Sender wie Empfänger müssen mindestens zwei Striche anzeigen.
- Ist die „Sensitivity“ beim Sender auf -18dB eingestellt.
- Ist „AF Out“ beim Empfänger auf 0dB eingestellt.
- Empfang beim Empfänger - auch während der Aufnahme - überprüfen.

Den Ausgang der Empfangseinheit (3,5mm Klinke Stecker) links bei der Kamera einstecken (Eingang ist mit MIC beschriftet).

### Kamera überprüfen und einstellen

- Bildeinstellungen: HG1920
- Mikrofon auf Automatik
- Videoaufnahme
- Intelligenter Autofokus (iA) aktiv?
- Ausreichender Speicherplatz?
- Ausreichende Akku-Laufzeit?



### Allgemeines vor Vorlesungsbeginn

- Mit dem/der Projektteilnehmer/-in abstimmen welche Aufnahmeart er/sie bevorzugt:
  - Nahaufnahme der Dolmetscher/-innen
  - Aufnahme der Dolmetscher/-innen und des/der Vortragenden
- Standort der Dolmetschern/-innen abklären. Falls der/die Studierende eine Aufnahme von Dolmetscher/-innen und Vortragenden/-r wünscht, die Position der Dolmetscher/-innen mit der des/der Vortragenden abstimmen – je geringer die Entfernung von Vortragendem/-r und Dolmetscher/-in ist, umso näher kann gezoomt werden.
- Die Dolmetscher/-innen sollten sich vor einem einfarbigen Hintergrund befinden (sofern der beeinflussbar und möglich ist).
- Dem/der Vortragendem/-n das Mikrofon befestigen - Brust Mitte, an nicht allzu beweglichen Kleidungsstücken (z.B.: am Sakko, der Hemdtasche,...) – und die Sendeeinheit übergeben (und Mute ausschalten).
- Auf Aufzeichnungsbedingungen hinweisen - Meldungen der Studierenden sollen vom/von der Vortragenden wiederholt werden. Ihn/sie bitten, nicht zu sehr umherzugehen. Eventuelle Fragen des/der Vortragenden klären.
- Aufnahmebereich einstellen
- Aufnahme starten – mit dem Record Knopf (der große länglichen Knopf mit dem rotem Punkt in der Mitte)

**Während der Aufnahme**

Je nach den Wünschen des/der anwesenden hörbehinderten Studierenden die Aufnahmeziele mit der Kamera erfassen. Falls er/sie es wünscht, das Geschehnis (Experimente, Videos,...) der Vorlesung im Blickfeld der Kamera behalten. Das Zielpublikum der Aufzeichnung sind zwar alle Studierenden, es wird jedoch auf die hörbehinderten/gehörlosen Studierenden besondere Rücksicht genommen. Daher auf die visuellen Informationen achten!

Aufnahmepriorität – sofern nicht anders vom/von der Projektteilnehmer/-in angegeben:

1. Dolmetscher/-innen (wenn vorhanden)
2. Tafel
3. Folien
4. Vortragende/-r

**Nach Ende der Vorlesung**

Equipment wieder abbauen.

### **3. Aufzeichnung mit einer im Hörsaal installierten LectureTube - Aufzeichnungseinheit der TU Wien**

#### **Vor Vorlesungsbeginn**

##### **Suche eines geeigneten Platzes**

- Mindestabstand vom Filmobjekt: 2m
- Abstand zur Aufzeichnungseinheit beachten (im Aufnahmeset ist für die Übertragung des Kamerasignales ein 10m Kabel vorhanden).
- Wenn Dolmetscher/-innen in der Vorlesung anwesend sind, einen Platz in der 1. - 3. Reihe wählen. Die Kamera sollte idealerweise auf Augenhöhe platziert werden können.
- Bei Aufnahmen ohne Dolmetscher/-innen Wahl eines Platzes, sodass genug Abstand zu den Folien und dem/der Vortragenden besteht.
- Bei Hörsälen ohne oder mit flach steigenden Reihen am besten einen Platz in der erste Reihe wählen.

##### **Kamera & Stativ vorbereiten**

Bei der Aufzeichnung der Gebärdensprachdolmetscher/-innen mit einer fest im Hörsaal installierten Aufzeichnungseinheit der TU Wien wird das Funkmikrofonset nicht verwendet. Der gesprochene Vortrag wird von der LectureTube-Einheit gemeinsam mit den Vorlesungsfolien aufgezeichnet.

##### **Kamera überprüfen und einstellen**

- Bildeinstellungen: HG1920
- Videoaufnahme
- Intelligenter Autofokus (iA) aktiv?
- Ausreichende Akku-Laufzeit?
- Ext. Display auf „Aus“?\*

\* Die Einstellung für das „Ext. Display“ ist im Menü unter „Einrichtung“ auf der zweiten Seite zu finden. Wichtig: Diese Einstellung wird beim Ausschalten der Kamera zurückgesetzt!

##### **Kamera und Aufzeichnungseinheit verbinden**

Die beigelegte Kabel-Peitsche am Stecker „AV-Multi“ verbinden. Er befindet sich an der linken Seite der Kamera. Auf dem Bildschirm erscheint nun ein Dialog. Hier „AV OUT“ auswählen. Den gelben Stecker der Kabel-Peitsche mit dem 10m Cinch-Verlängerungskabels mit der Aufzeichnungseinheit verbinden.

### **Allgemeines vor Vorlesungsbeginn**

- Mit dem/der Projektteilnehmer/-in abstimmen welche Aufnahmeart er/sie bevorzugt:
  - Nahaufnahme der Dolmetscher/-innen
  - Aufnahme der Dolmetscher/-innen und des/der Vortragenden
- Standort der Dolmetschern/-innen abklären. Falls der/die Studierende eine Aufnahme von Dolmetscher/-innen und Vortragenden/-r wünscht, die Position der Dolmetscher/-innen mit der des/der Vortragenden abstimmen – je geringer die Entfernung von Vortragendem/-r und Dolmetscher/-in ist, umso näher kann gezoomt werden.
- Die Dolmetscher/-innen sollten sich vor einem einfarbigen Hintergrund befinden (sofern der beeinflussbar und möglich ist).
- Dem/der Vortragendem/-n das Mikrofon befestigen - Brust Mitte, an nicht allzu beweglichen Kleidungsstücken (z.B.: am Sakko, der Hemdtasche,...) – und die Sendeeinheit übergeben (und Mute ausschalten).
- Auf Aufzeichnungsbedingungen hinweisen - Meldungen der Studierenden sollen vom Vortragenden wiederholt werden. Ihn/sie bitten nicht zu sehr umherzugehen. Eventuelle Fragen des/der Vortragenden klären.
- Aufnahmebereich einstellen.
- Die Lecture-Tube Einheit startet die Aufzeichnung automatisch. Die Vorlesung dennoch mit der Videokamera aufzeichnen, falls die Aufzeichnung durch die Lecture-Tube-Einheit nicht funktioniert. Die Aufnahme kann gelöscht werden, sobald die einwandfreie Aufzeichnung online ist.

### **Während der Aufnahme**

Je nach den Wünschen des/der anwesenden hörbehinderten Studierenden die Aufnahmeziele mit der Kamera erfassen. Falls er/sie es wünscht, das Geschehnis (Experimente, Videos,...) der Vorlesung im Blickfeld der Kamera behalten. Das Zielpublikum der Aufzeichnung sind zwar alle Studierenden, es wird jedoch auf die hörbehinderten/gehörlosen Studierenden besondere Rücksicht genommen. Daher auf die visuellen Informationen achten!

Aufnahmepriorität – sofern nicht anders vom/von der Projektteilnehmer/-in angegeben:

1. Dolmetscher/-innen (wenn vorhanden)
2. Tafel
3. Folien
4. Vortragende/-r

### **Nach Ende der Vorlesung**

Equipment wieder abbauen.

### 9.2.3 eTutor Leitfaden für die Nachbearbeitung und das Onlinestellen von Lehrveranstaltungsaufzeichnungen



#### Lehrveranstaltungsaufnahme Videobearbeitung

Der Leitfaden dient zur Übersicht und als Hilfestellung für die Nachbearbeitung der aufgenommen Videos und dem Hochladen der Videos auf die Streaming-Plattform der jeweiligen Universität.

Bei Fragen Werner Nemecek unter 01/234 567 kontaktieren.

#### Inhalt

|   |          |
|---|----------|
| <b>1 Video-Nachbearbeitung .....</b>                                    | <b>2</b> |
| <b>1.1 Nachbearbeitung unter Windows .....</b>                          | <b>2</b> |
| 1.1.1 Videoschnitt .....  | 3        |
| 1.1.2 Videokonvertierung .....  | 3        |
| 1.1.3 Installation und Konfiguration von TSSnipper .....                | 5        |
| 1.1.4 Installation und Konfiguration von Any Video Converter .....      | 5        |
| <b>1.2 Nachbearbeitung unter Mac OS X .....</b>                         | <b>6</b> |
| 1.2.1 Turbo.264 HD Einstellungen und Profile .....                      | 8        |
| <b>2 Onlinestellen .....</b>  | <b>9</b> |
| <b>2.1 Universität Wien.....</b>  | <b>9</b> |
| <b>2.2 TU Wien.....</b>   | <b>9</b> |
| 2.2.1 Onlinestellen von Videos bei der automatischen Aufzeichnung ..... | 9        |
| 2.2.2 Onlinestellen von Videos bei manueller Aufzeichnung .....         | 9        |

## 1 Video-Nachbearbeitung

Bevor die Videos im Streamingservice der jeweiligen Universität zur Verfügung gestellt werden können, müssen die Aufnahmen nachbearbeitet werden. Bei der Aufzeichnung verteilt die Kamera in der Regel die Videos auf mehrere Dateien. Die einzelnen Files müssen zusammengefügt, eventuell geschnitten und in ein anderes Format konvertiert werden.

Falls erforderlich, muss der Beginn sowie das Ende des Videos weggeschnitten werden. Es kommt auch vor, dass der/die Vortragende einen Teil der Vorlesung nicht aufgezeichnet haben möchte. Die betreffenden Abschnitte müssen aus dem Video herausgeschnitten werden.

Die genaue Anleitung zum Schnitt und zur Konvertierung findet sich in den Abschnitten zu den jeweiligen Betriebssystemen. Dabei handelt es sich um Softwareempfehlungen. Andere Programme können ebenfalls verwendet werden, sofern ein gleichwertiges Ergebnis erzielt wird.

Folgende Dateinamen werden für die fertigen Dateien verwendet:

- Universität Wien: Vorlesungsnummer\_JJMMTT (123456\_120526)
- TU Wien: Vorlesungsnummer\_JJMMTT (123456\_120526)

### 1.1 Nachbearbeitung unter Windows

Für Windows werden die Programme TSSnipper zum Schneiden und Any Video Converter zum Konvertieren und Zusammenfügen der einzelnen Videos verwendet. Hinweise zur Installation und Konfiguration der beiden Programme sind am Ende des Abschnittes angeführt.

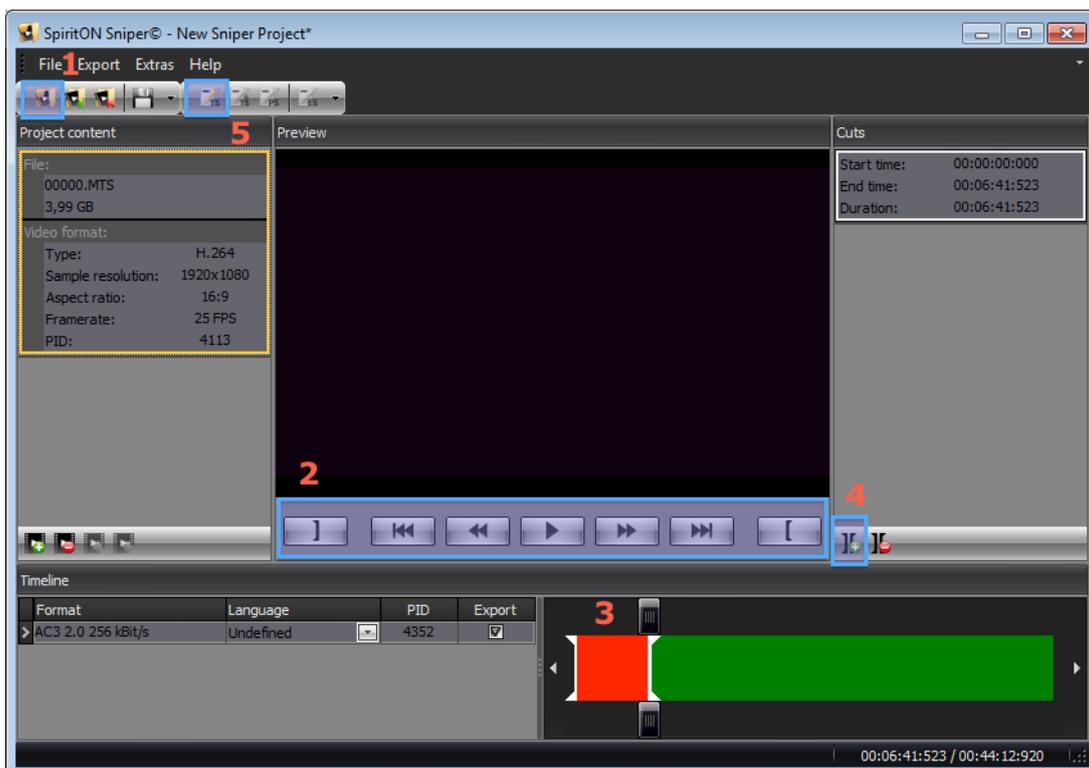
### 1.1.1 Videoschnitt

Mit TSSniper werden die unbearbeiteten Originaldateien der Kamera (\*.mts) geschnitten.

Im Programm muss zuerst ein neues Projekt über den Button „Create a new Project“ (im Screenshot #1) geöffnet werden. Das zu schneidende Video über den Menüpunkt „File“ > „Add a new file to the project“ hinzufügen (dies ist auch mittels Rechtsklick auf die Fläche „Project content“ möglich). Im Dateialog müssen alle Dateien zur Anzeige ausgewählt werden. Standardmäßig werden nur Dateien mit der Endung \*.ts angezeigt.

Für das Video können durch die Steuerelemente (im Screenshot #2) Schnittbereiche definiert werden. Der Button „]“ markiert den Beginn eines Schnittbereichs, der Button „[“ das Ende. In der Zeitleiste des Programms (im Screenshot #3) werden die Bereiche rot markiert, die aus dem Video geschnitten werden sollen. Wurde der zu entfernende Bereich festgelegt, muss er mit dem Button „add a new area from the current selection“ (im Screenshot #4) bestätigt werden.

Sind alle zu entfernenden Abschnitte definiert kann das Video mit dem Button „export transport stream“ (im Screenshot #5) gespeichert werden. Vor dem Sichern noch die Dateiendung auf \*.mts ändern.

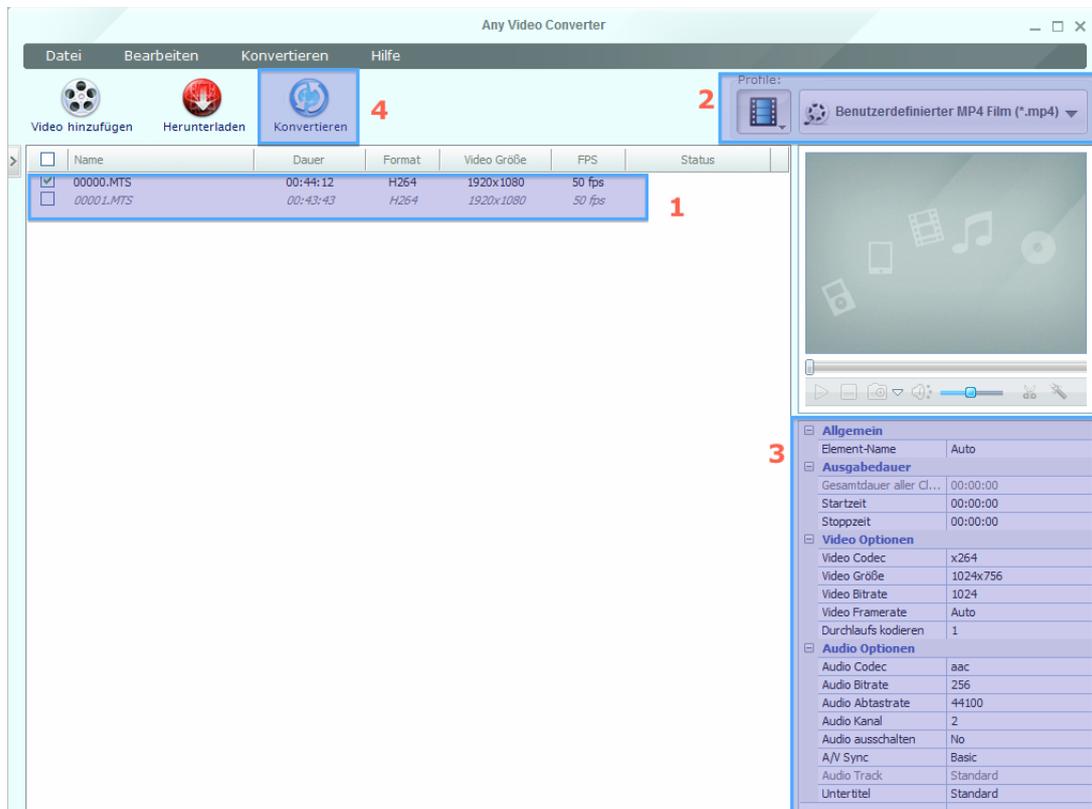


### 1.1.2 Videokonvertierung

Mit dem Programm Any Video Converter werden die einzelnen Videodateien zusammengefügt und in das passende Format umgewandelt.

Die zu konvertierenden Videos sind in den Hauptbereich des Programms zu ziehen. Danach erscheinen deren Namen sowie Informationen zu den Videodateien. Es sind sämtliche Dateien zu markieren, die zusammengefügt werden sollen. Durch einen Klick der rechten Maustaste auf die Auswahl das Kontextmenü aufrufen. Dort den Punkt „Auswahl zusammenführen“ auswählen. Die zusammengeführten Videos erkennt man an

der grauen Schrift. Die Reihenfolge der Videos ist immer von oben nach unten. D.h. das erste Video bildet den Anfang, das letzte das Ende.



Wurden die oben angeführten Schritte ausgeführt, muss noch das Ausgabeformat des Videos eingestellt werden. Falls nicht bereits automatisch von AVC ausgewählt, rechts oben „Benutzerdefinierter MP4 Film (\*.mp4)“ auswählen (siehe im Screenshot #2). Im Anschluss daran müssen noch die Konvertierungsoptionen eingestellt werden. Sie können im Bereich rechts unten eingestellt werden (siehe im Screenshot Bereich #3):

- Video Codec: x264
- Video Größe: 1024 x 576
- Video Bitrate: 1024
- Video Framerate: Auto
- Audio Codec: AAC
- Audio Bitrate: 256
- Rest: Automatisch bzw. die Voreinstellungen verwenden

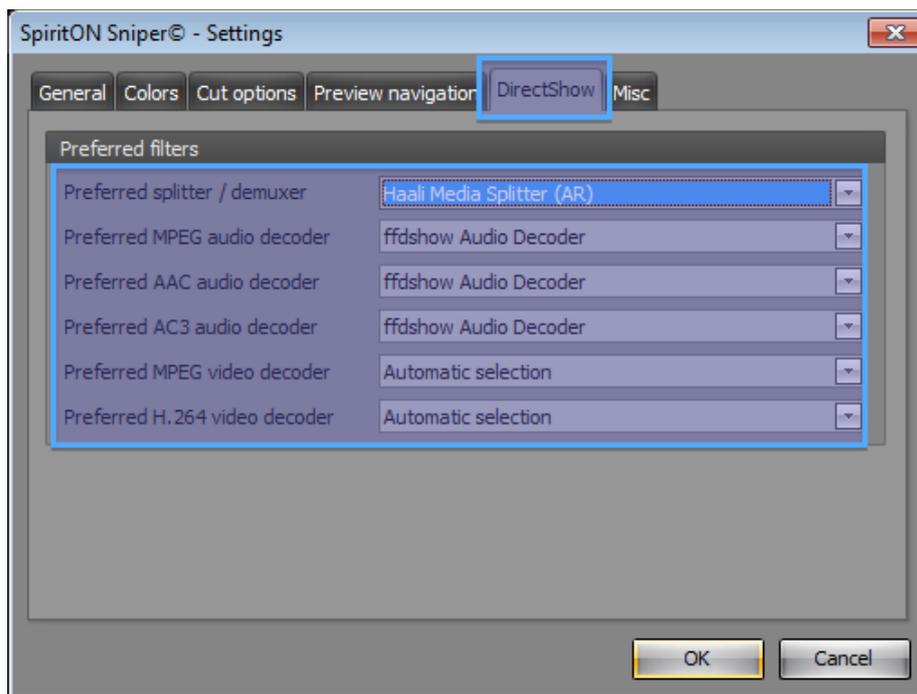
Sobald alles eingestellt ist kann die Konvertierung mittels des Buttons „Konvertieren“ (im Screenshot #4) gestartet werden.

### 1.1.3 Installation und Konfiguration von TSSniper

Damit TSSniper funktioniert, müssen folgende Dateien heruntergeladen und installiert werden. Es handelt sich um das Programm selbst sowie zwei Video-Codecs:

- <http://www.videohelp.com/tools/TSSniper>
- <http://www.videohelp.com/tools/ffdshow>
- <http://www.videohelp.com/tools/Haali-Media-Splitter>

Im Anschluss an die Installation muss in den Einstellungen von TSSniper im Tab „DirectShow“ bei „Preferred splitter / demuxer“ „Haali Media Splitter (AR)“ und bei den drei „Audio Decoder“ Feldern „ffdshow Audio Decoder“ ausgewählt werden:



### 1.1.4 Installation und Konfiguration von Any Video Converter

Any Video Converter kann auf der folgenden Seite heruntergeladen werden:

[http://www.any-video-converter.com/products/for\\_video\\_free/](http://www.any-video-converter.com/products/for_video_free/)

## 1.2 Nachbearbeitung unter Mac OS X

Für Mac OS X wird Turbo.264 HD zur Nachbearbeitung der Videos verwendet. Das Programm wird nach Rücksprache von GESTU verliehen.

Die Videoteile sind gemeinsam in das Turbo.264 HD Fenster zu ziehen. Die verschiedenen Teile der Aufnahme sollten dabei als ein Eintrag erscheinen:



Der Titel des Ausgabevideos wird durch einen Klick auf den Namen der Originaldatei (im Screenshot #2) geändert. Das Profil beziehungsweise die Einstellungen für die Konvertierung des Videos werden über die Dropdown Liste „Format“ (im Screenshot #3) ausgewählt. (Näheres zu den Profilen im nächsten Kapitel - 1.2.1 Turbo.264 HD Einstellungen und Profile).

Achtung: im Hauptfenster darf „Breitbild“ nicht markiert werden!

Mit einem Klick auf das Vorschaubild wird der Schnittditor des Programms geöffnet (Screenshot #1):

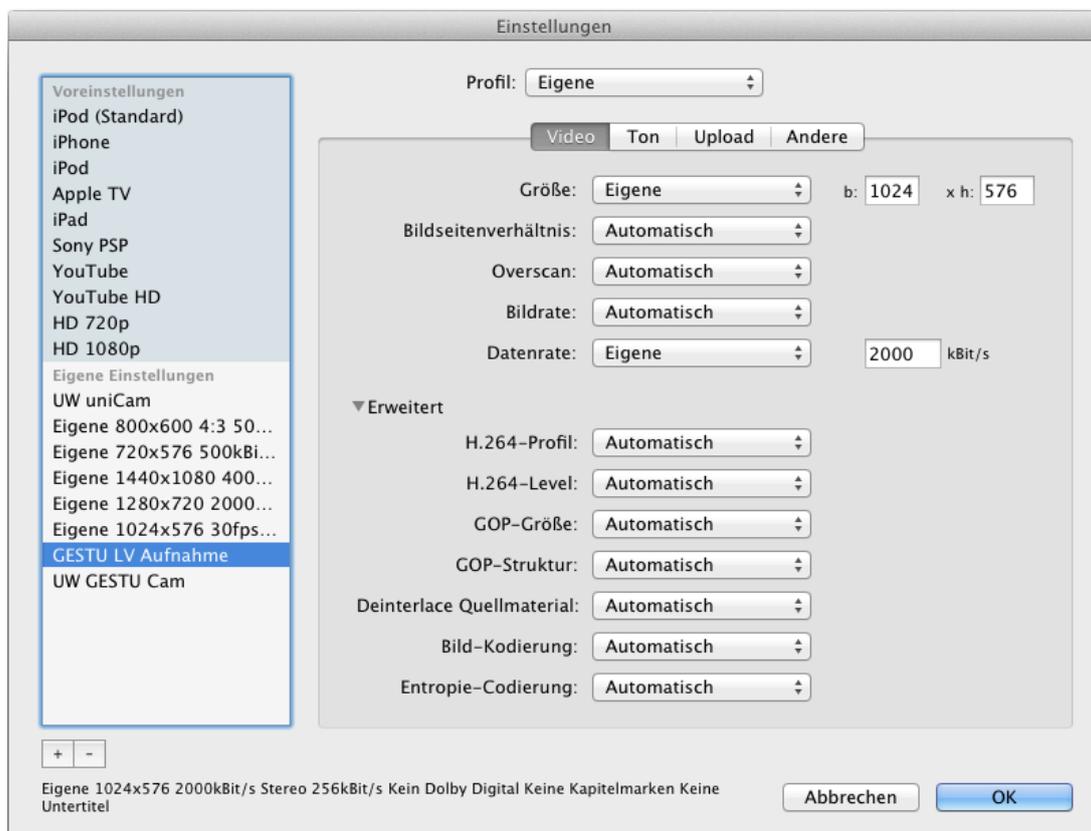


Im Schnittfenster das gewünschte Video auswählen (im Screenshot #1). Die Schnittmarken können durch das betätigen des, im Screenshot mit der #2 markierten, Buttons eingestellt werden.

### 1.2.1 Turbo.264 HD Einstellungen und Profile

Mit dem Programm Turbo.264 HD können Profile für die Konvertierung angelegt werden. Die Profile können in der Dropdownliste „Format“ im Turbo.264 HD Hauptfenster, wenn sich Videos im Programm befinden, durch klicken des Punktes „Bearbeiten“ editiert oder neu erstellt werden. Dabei die folgenden Parameter angeben:

- Größe: 1024 x 576 Pixel
- Datenrate: Eigene - 2000 kBit/s
- Ton: Stereo, 256kBit/s
- Rest: Automatisch



## 2 Onlinestellen

### 2.1 Universität Wien

Im Streamingsupport-Center der Uni Wien auf einem der dort bereitgestellten Computer mit dem Explorer zu folgendem Share verbinden:

\\hier war die Adresse zum Share

Anmeldung mit UniVie Benutzernamen und Passwort.

Video in den entsprechenden Ordner kopieren: Speicherort/der/Dateien

Alternativ kann der Upload auch von zu Hause via VPN erfolgen. Die VPN Einstellungen sind hier zu finden: <http://VPNEinstellungen.com>

Sobald eine Verbindung zur Universität Wien via VPN hergestellt ist mit dem zuvor genannten Server verbinden und das Video hochladen.

Sobald das Video im Streaming-service zum Abspielen zur Verfügung steht, erhält man ein E-Mail vom Uni-Wien Streamingserver. Das Mail bitte an mich ([nemecek@is.tuwien.ac.at](mailto:nemecek@is.tuwien.ac.at)) weiterleiten. Den Link zur Aufnahme der Vorlesung dem/der GESTU-Teilnehmer/-in senden sowie das Video im eLearning-Kurs der Vorlesung einbinden, sofern der/die Vortragende dem zugestimmt hat. Auf Wunsch des/der Vortragenden oder des/der Dolmetschers/-in, sind ihnen die Links zu den Aufzeichnungen zu übersenden.

### 2.2 TU Wien

#### 2.2.1 Onlinestellen von Videos bei der automatischen Aufzeichnung

Bei der automatischen Aufzeichnung mit einer LectureTube Einheit wird der Link zur Aufzeichnung vom Teaching Support Center der TU Wien via E-Mail zugesendet.

#### 2.2.2 Onlinestellen von Videos bei manueller Aufzeichnung

Die fertig bearbeiteten Videos auf dem FTP Server des GESTU Projekts in den Ordner hochladen.

- Die Adresse des Servers lautet: Server Adresse
- Port: Port Nummer
- Verbindung über FTP

Der Benutzername und das Passwort wird von mir bekanntgegeben.

Nach dem Upload Herrn Mustermann vom Teaching Support Center der TU Wien via E-Mail informieren, dass das Video am Server hochgeladen wurde. Er stellt das Video auf der Streaming-Plattform Matterhorn zur Verfügung und sendet ein E-Mail mit dem Link zum Video zurück.

Danach das Video vom GESTU FTP Server wieder löschen.

# Abbildungsverzeichnis

|     |   |    |
|-----|---|----|
| 1.1 | Zuordnung und Abgrenzung dieser Diplomarbeit im GESTU Projekt sowie der beiden anderen im Zuge des Projekts durchgeführten Diplomarbeiten von Hattinger [Hat11] und [Hat13]. . . . .                            | 5  |
| 2.1 | Verschiedene Methoden zur manuellen Untertitel-Erzeugung [Hat11, Abbildungen 3.7 und 3.9]. . . . .  | 13 |
| 3.1 | Einflussfaktoren auf ein Sprachsignal [PK08, Abbildung 2.1]. . . . .  | 19 |
| 3.2 | Schematischer Ablauf einer Spracherkennung [RF11, Figure 5.1]. . . . .  | 22 |
| 5.1 | Das Prinzip von Synote [WWM <sup>+</sup> 09, Figure 1]. . . . .   | 58 |
| 5.2 | Das Synote Benutzerinterface [WWM <sup>+</sup> 09, Figure 3]. . . . .   | 58 |
| 5.3 | Zwei der gebräuchlichsten Bildeinstellungen für die Videoaufzeichnung von Lehrveranstaltungen für hörbehinderte Studierende, die im Zuge des GESTU Projekts durchgeführt wurden. . . . .                        | 62 |
| 6.1 | Der Eintrag für „Bär“ in einem ASL Lexikon [Zwi10, Figure 2]. . . . .   | 76 |
| 6.2 | Die ÖGS Gebärde für „GESTU“, dargestellt durch zwei Bilder [Hat11, Abbildung 1.2]. . . . .  | 76 |
| 6.3 | Die Darstellung der ASL Gebärde für „Bär“ in drei verschiedenen Notationssystemen [Zwi10, Figure 4]. . . . .  | 76 |
| 6.4 | Suchmaske zur Auswahl des Ausführungsortes der Gebärde in einem online Gebärdenlexikon [Zwi10, Figure 8]. . . . .   | 78 |
| 6.5 | Der Eintrag für „Array“ im ASL-STEM Forum [Cav10, Figure 5.2]. . . . .  | 82 |
| 6.6 | Der Eintrag für den Begriff „Exponent“ im GESTU Fachgebärden Wiki. . . . .  | 85 |
| 6.7 | Ein Ausschnitt aus einem, von einem Benutzer des Fachgebärden Wikis selbstgestellten qualitativ hochwertigen Video. . . . .   | 90 |
| 6.8 | Ausschnitte aus zwei, von Benutzerinnen des GESTU Fachgebärden Wikis, selbstgestellten Videos. Bei der Aufnahme oder Nachbearbeitung wurden nicht alle auf Seite 88 genannten Empfehlungen eingehalten. . . . . | 91 |
| 7.1 | RGD bei zwei möglichen Szenarien [Hat11, Abbildung 3.12]. . . . .   | 94 |
| 7.2 | Eine schematische Darstellung des Prinzips von ClassInFocus [Cav10, Figure 4.1]. . . . .  | 95 |
| 7.3 | ClassInFocus im Einsatz [Cav10, Figure 4.2]. . . . .  | 96 |

# Tabellenverzeichnis

|     |   |    |
|-----|---|----|
| 2.1 | Vergleich manueller live Untertitelungsmethoden . . . . . | 14 |
|-----|---|----|



# Literaturverzeichnis

- [Aig12] AIGNER, Theresa: *Studiendauer für Gehörlose: 20 Jahre*. [http://diepresse.com/home/bildung/universitaet/1267836/Studiendauer-fuer-Gehoerlose\\_20-Jahre](http://diepresse.com/home/bildung/universitaet/1267836/Studiendauer-fuer-Gehoerlose_20-Jahre). Version: Juli 2012
- [BA04] BROTHERTON, Jason A. ; ABOWD, Gregory D.: Lessons learned from eClass: Assessing automated capture and access in the classroom. In: *ACM Transactions on Computer-Human Interaction* 11 (2004), Nr. 2. <http://dl.acm.org/citation.cfm?id=1005362>
- [BOD<sup>+</sup>08] BIGHAM, Jeffrey P. ; OTERO, Daniel ; DEWITT, Jessica N. ; CAVENDER, Anna C. ; LADNER, Richard E.: ASL-STEM Forum: A Bottom-Up Approach to Enabling American Sign Language to Grow in STEM Fields. In: *International Conference on Weblogs and Social Media*, 2008
- [Bun11] BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND TECHNOLOGIE: *Nutzung des Frequenzbereichs 470 – 862 MHz durch Funk-Mikrofone (PMSE)*. 2011
- [Bun12] BUNDESMINISTERIUM FÜR WISSENSCHAFT UND FORSCHUNG: *Bestmögliche Rahmenbedingungen für gehörlose Studierende - GESTU wird fortgesetzt*. [http://www.ots.at/presseaussendung/OTS\\_20120712\\_OTS0112](http://www.ots.at/presseaussendung/OTS_20120712_OTS0112). Version: Juli 2012
- [Cav10] CAVENDER, Anna C.: *Collaborative, multimedia solutions for improving educational access for deaf and hard of hearing students*, University of Washington, Dissertation, 2010. <http://gradworks.umi.com/34/21/3421543.html>
- [Cha07] CHANG, Shanton: Academic perceptions of the use of Lectopia: A University of Melbourne example. In: *Proceedings ascilite Singapore 2007*, 2007
- [DCL09] DAVIS, Simon ; CONNOLLY, Anthea ; LINFIELD, Edmund: Lecture capture: Making the most of face to face learning. In: *Engineering Education* 4 (2009), Nr. 2. <http://84.22.166.132/journal/index.php/ee/article/viewArticle/132/170.html>

- [Echa] ECHO360: *Gallaudet University — Liberal Arts Program for the Deaf and Hearing-Impaired*. <http://echo360.com/customers/case-studies/gallaudet/>. – zuletzt abgerufen am 07.03.2013
- [Echb] ECHO360: *Lectopia*. <http://www.lectopia.com.au/>. – zuletzt abgerufen am 29.01.2013
- [FHK08] FRIEDLAND, Gerald ; HÜRST, Wolfgang ; KNIPPING, Lars: Automated Lecture Recording. In: FURHT, Borko (Hrsg.): *Encyclopedia of multimedia*. Springer Verlag, 2008. – ISBN 0387747249
- [Hat11] HATTINGER, Christian F.: *GESTU: Evaluierung von technischen Hilfsmitteln zur Förderung Studierender mit Hörbehinderung im österreichischen tertiären Bildungssektor und Einführung geeigneter Technologien an der TU Wien*, Technische Universität Wien, Diplomarbeit, 2011
- [Hat13] HATTINGER, Christian F.: *GESTU: Erarbeitung eines Respeaking Trainingsprogramms und Vergleich zur automatischen Spracherkennung, um diese Techniken zur Untertitelerzeugung in E-Learning Plattformen zu verwenden und somit speziell hörbeeinträchtigte Studierende zu unterstützen*, Technische Universität Wien, Diplomarbeit, 2013
- [Her11] HERMANN, Christoph: *Techniken und Konzepte für den nachhaltigen Einsatz von Vorlesungsaufzeichnungen*, Universität Freiburg, Dissertation, 2011. <http://www.freidok.uni-freiburg.de/volltexte/8699/>
- [Hic13] HICKERSBERGER, Helmut: *Neue Algorithmen der Automatischen Spracherkennung*, Technische Universität Wien, Dissertation, 2013
- [KS07] KRAUSNEKER, Verena ; SCHALBER, Katharina: *Sprache Macht Wissen / Innovationszentrum der Universität Wien; Verein Österreichisches Sprachen-Kompetenz-Zentrum*. 2007. – Forschungsbericht
- [Lar10] LARKIN, Helen E.: But they won't come to lectures..."The impact of audio recorded lectures on student experience and attendance. In: *Australasian journal of educational technology* 26 (2010), Nr. 2. <http://ascilite.org.au/ajet/ajet26/larkin.html>
- [McC08] MCCCLURE, Ann: *Lecture Capture: A Fresh Look*. In: *University Business* (2008). <http://www.universitybusiness.com/article/lecture-capture-fresh-look>
- [Mer07] MERTENS, Robert: *Hypermediale Navigation in Vorlesungsaufzeichnungen*, Universität Osnabrück, Dissertation, 2007
- [MPC<sup>+</sup>05] MARSCHARK, Marc ; PELZ, Jeff B. ; CONVERTINO, Carol ; SAPERE, Patricia ; ARNDT, Mary E. ; SEEWAGEN, Rosemarie: *Classroom Interpreting and Visual*

- Information Processing in Mainstream Education for Deaf Students: Live or Memorax(R)? In: *American Educational Research Journal* 42 (2005), Nr. 4. <http://aer.sagepub.com/cgi/doi/10.3102/00028312042004727>
- [MPGH12] MÜLLER, Meinolf ; POLLMÄCHER, Dirk ; GERTH, Michael ; HAASE, Kristina: *Vorlesungsaufzeichnung*. 2012
- [Net07] NET4VOICE: Net4Voice - Document on Methodology Specification. 2007. – Forschungsbericht
- [ORF12] ORF: 3. *Etappenplan zum Ausbau des Barrierefreien Zugangs zu den ORF-Fernseh-programmen und zum ORF-Online-Angebot gemäß§3 Abs. 1Z2 ORF-Gesetz*. [http://zukunft.orf.at/show\\_content.php?sid=78&pvi\\_id=1187](http://zukunft.orf.at/show_content.php?sid=78&pvi_id=1187). Version: 2012
- [PK08] PFISTER, Beat ; KAUFMANN, Tobias: *Sprachverarbeitung - Grundlagen und Methoden der Sprachsynthese und Spracherkennung*. 2008. – ISBN 9783540759096
- [RF11] ROMERO-FRESCO, Pablo: *Subtitling through Speech Recognition: Respeaking*. St Jerome Publishing, 2011. – ISBN 190576328X
- [RFM13] ROMERO-FRESCO, Pablo ; MARTÍNEZ, Juan: *Accuracy rate in live subtitling – the NER model*. 2013
- [RLS11] RECTOR, Kyle ; LADNER, Richard ; SHEPARDSON, Michelle: Incentivizing the ASL-STEM forum. In: *Proceedings of the 7th International Symposium on Wikis and Open Collaboration, 2011 (WikiSym '11)*
- [Roc09] ROCKWELL, Mark: Cameras Enhance Teaching and Learning at Gallaudet University. In: *EdTech Magazine* (2009). <http://www.edtechmagazine.com/higher/article/2009/02/visual-cues>
- [SAH10] STEWART, Iain ; ALLAN, Malcolm ; HARRISON, David: The development of a lecture capture system based on a tool to support hearing impaired students. In: *Technological Developments in Education and Automation*. 2010
- [SBG10] SECKER, Jane ; BOND, Steve ; GRUSSENDORF, Sonja: Lecture capture: rich and strange, or a dark art? In: *17th International Conference of the Association for Learning Technology, 2010*
- [Sta] STATISTIK AUSTRIA: *Statistiken zum Studium*. [http://www.statistik.at/web\\_de/statistiken/bildung\\_und\\_kultur/formales\\_bildungswesen/universitaeten\\_studium/index.html](http://www.statistik.at/web_de/statistiken/bildung_und_kultur/formales_bildungswesen/universitaeten_studium/index.html). – zuletzt abgerufen am 24.10.2012
- [Sta12] STATISTIK AUSTRIA: *Bildung in Zahlen 2010/11*. [http://www.statistik.at/web\\_de/services/publikationen/5/index.html?id=5&listid=5&detail=508](http://www.statistik.at/web_de/services/publikationen/5/index.html?id=5&listid=5&detail=508). Version: 2012

- [Tal12] TALBERT, Robert: Inverted Classroom. In: *Colleagues* 9 (2012), Nr. 1. <http://scholarworks.gvsu.edu/colleagues/vol9/iss1/7/>
- [TGN<sup>+</sup>10] TIBALDI, Daniela ; GARLASCHELLI, Luca ; NARDONE, Mariarosaria ; SANTINI, Marco ; RIGONI, Matteo ; WALD, Mike ; HOWARD, Chris ; WEGGERLE, Alexander ; SCHULTHESS, Peter: Net4Voice Final Report. Version: 2010. [https://www.net4voice.eu/net4voice/Deliverables/Net4Voice\\_Final\\_Report.pdf](https://www.net4voice.eu/net4voice/Deliverables/Net4Voice_Final_Report.pdf). 2010. – Forschungsbericht
- [Top11] TOPPIN, Ian N.: Video lecture capture (VLC) system: A comparison of student versus faculty perceptions. In: *Education and Information Technologies* 16 (2011), Nr. 4. <http://www.springerlink.com/index/10.1007/s10639-010-9140-x>
- [Wal06] WALD, Mike: An exploration of the potential of Automatic Speech Recognition to assist and enable receptive communication in higher education. In: *ALT-J: Research in Learning Technology* 14 (2006), Nr. 1. <http://repository.alt.ac.uk/108/>
- [Wal10] WALD, Mike: Synote: Accessible and Assistive Technology Enhancing Learning for All Students. In: *ICCHP 2010, Part II, LNCS 6180* (2010). <http://eprints.soton.ac.uk/271147/>
- [WHO05] WHO: *ICF - Die Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit*. 2005 <http://www.dimdi.de/static/de/klassi/icf/index.htm>
- [Wil06] WILLIAMS, Jocasta: The Lectopia service and students with disabilities. In: *Proceedings of the 23rd annual ascilite conference: Who's learning? Whose technology?*, 2006
- [WL12] WALD, Mike ; LI, Yunjia: Synote: Important Enhancements to Learning with Recorded Lectures. In: *2012 IEEE 12th International Conference on Advanced Learning Technologies* (2012), Juli. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6268168>
- [WWM<sup>+</sup>09] WALD, Mike ; WILLS, Gary ; MILLARD, David ; GILBERT, Lester ; KHOJA, Sha-keel ; KAJABA, Jiri ; LI, Yunjia ; SINGH, Priyanaka: ENHANCING LEARNING USING SYNCHRONISED MULTIMEDIA ANNOTATION. In: *EUNIS 2009: - IT: Key of the European Space for Knowledge*, 2009
- [You12] YOUTUBE: *YouTube automatic captions now in six European languages*. <http://youtube-global.blogspot.de/2012/11/youtube-automatic-captions-now-in-six.html>. Version: 2012. – zuletzt abgerufen am 11.03.2013

- [Zwi10] ZWITSERLOOD, Inge: Sign Language Lexicography in the Early 21st Century and a Recently Published Dictionary of Sign Language of the Netherlands. In: *International Journal of Lexicography* 23 (2010), Nr. 4. <http://ijl.oxfordjournals.org/cgi/doi/10.1093/ijl/ecq031>



# Glossary

- ASL** American Sign Language - im Deutschen: amerikanische Gebärdensprache
- ASR** Automatic Speech Recognition - im Deutschen: automatische Spracherkennung
- DNS** Dragon NaturallySpeaking ist eine Spracherkennungssoftware der Firma Nuance Communications
- EML** European Media Laboratory
- FPS** Frames per Second - im Deutschen: Bilder pro Sekunde
- GESTU** Gehörlos Erfolgreich Studieren an der TU Wien
- ICF** International Classification of Functioning, Disability and Health - im Deutschen: Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit
- LLC** Liberated Learning Consortium
- MHz** Mega Hertz - ist die SI-Einheit für die Frequenz
- MP4** MPEG-4 Part 14 ist ein Dateiformat für Video-, Audio- oder Textdateien
- MySQL** Ein Datenbanksystem
- NER** Ein Modell zur Bestimmung der Qualität von Untertiteln. Es berücksichtigt unter anderem die Auswirkung der Fehler und Umformulierungen.
- Net4Voice** War ein Teilprojekt des europäischen „Lifelong Learning“ Programms
- ÖGS** Österreichische Gebärdensprache
- PDF** Portable Document Format - ist ein Dateiformat für Dokumente
- PHP** PHP: Hypertext Preprocessor ist eine Scriptsprache mit der Webanwendungen erstellt werden können

**RGD** Remote **G**ebärdensprach**d**olmetschen

**Ruby on Rails** Ist ein Web Application Framework, geschrieben in der Programmiersprache Ruby

**STEM** Dabei handelt es sich um die Anfangsbuchstaben für die englischen Fachbereiche **S**cience (Naturwissenschaft), **T**echnology (Technik), **E**ngineering (Ingenieurwissenschaft), und **M**athmatics (Mathematik)

**WER** **W**ord **E**rror **R**ate - im Deutschen: Wortfehlerrate. Sie ist ein Qualitätsmerkmal für Spracherkennungssysteme

**WHO** **W**orld **H**ealth **O**rganization - im Deutschen: Weltgesundheitsorganisation

**WLAN** **W**ireless **L**ocal **A**rea **N**etwork

**WpM** **W**ords **p**er **M**inute - im Deutschen: Wörter pro Minute

**WRR** **W**ord **R**ecognition **R**ate - im Deutschen: Wortakkuratheit. Sie ist ein Qualitätsmerkmal für Spracherkennungssysteme