MASTERARBEIT | MASTER'S THESIS

Titel | Title

Computergestütztes Dolmetschen - Die Anwendung von Sight-Terp beim Konsekutivdolmetschen im Sprachenpaar Englisch-Deutsch

verfasst von | submitted by

Katharina Schmidt BA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 070 331 345
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Translation Deutsch Französisch
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Franz Pöchhacker

Danksagung

Zuallererst möchte ich mich herzlich bei meinem Betreuer Herrn Univ.-Prof. Mag. Dr. Franz Pöchhacker für zahlreiche Verbesserungsvorschläge, die konstruktive Kritik und vor allem eine offene Diskussionsbereitschaft bedanken. Ihre wertvollen Tipps und Ratschläge waren mir eine große Hilfe.

Ich bedanke mich ebenso bei allen Wegbegleiter*innen und Freund*innen, die ich im Laufe meines Bachelor- und Masterstudiums am Zentrum für Translationswissenschaft kennenlernen durfte. Insbesondere weiß ich die Kooperationsbereitschaft der teilnehmenden Kolleg*innen am durchgeführten Experiment und Leitfaden-Interview zu schätzen. Die investierte Zeit und Energie haben spannende Einblicke in das computergestützte Konsekutivdolmetschen geliefert und das Entstehen dieser Arbeit überhaupt erst ermöglicht.

Abschließend bedanke ich mich bei meiner Familie, die mich auf meinem langen Weg durch das Studium auch in schwierigen Momenten unterstützt und ermutigt hat, mein Ziel nicht aus den Augen zu verlieren.

Inhaltsverzeichnis

Abbildungs- und Tabellenverzeichnis

Abkürzungsverzeichnis

AI	artificial intelligence
AIIC	Association Internationale des Interprètes de Conférence, internationaler Verband der Konferenzdolmetscher*innen
ALPAC	Automatic Language Processing Advisory Committee
ASR	automatic speech recognition, automatische Spracherkennung
CACI	computer-assisted consecutive interpreting, computergestütztes Konsekutivdolmetschen
CAI	computer-assisted interpreting, computergestütztes Dolmetschen
CAT	computer-assisted translation, computergestützte Übersetzung
ELF	Englisch als Lingua Franca
ICT	information and communication technologies, Informations- und Kommunikationstechnologien
ISO	International Organization for Standardization
KI	künstliche Intelligenz
MÜ	maschinelle Übersetzung
NE	Named Entity
NMT	neural machine translation, neuronale maschinelle Übersetzung
PE	post editing, postmaschinelle Bearbeitung
Pos.	Position
PSSH	Physical Symbol Systems Hypothesis
PT	problem trigger, Problemauslöser
RSI	remote simultaneous interpreting, simultanes Ferndolmetschen
S2T	Speech-to-Text
SE	Sinneinheit, unit of sense
TN	Teilnehmende
VIP	Voice-text Integrated system for interPreters
VRI	video remote interpreting, Video-Ferndolmetschen
WER	word error rate
wpm	words per minute, Wörter pro Minute

Einleitung

Das sogenannte computergestützte Dolmetschen (CAI – ‚computer-assisted interpreting') ist seit einigen Jahren weder aus der Berufs- noch aus der Forschungspraxis von Dolmetscher*innen wegzudenken. Fantinuoli (2018b: 4) spricht sogar von einer technologischen Wende (‚technological turn') und führt darunter drei Hauptbereiche an, nämlich das computergestützte Dolmetschen, das Ferndolmetschen und das maschinelle Dolmetschen.

Technologien formen also bereits jetzt die Zukunft des Dolmetschberufs, jedoch wird das Thema im akademischen Diskurs zu wenig angesprochen (Corpas Pastor, 2021b: 93). Fantinuoli (2016: 50) kritisiert ebenso, dass trotz eines gestiegenen Interesses am computergestützten Dolmetschen dieses in der Forschung bisher wenig Aufmerksamkeit erhielt. Aus diesem Grund soll sich die vorliegende Masterarbeit mit dem Thema des computergestützten Dolmetschens genauer beschäftigen.

Zunächst wird dafür eine Unterteilung getroffen. CAI-Tools können zur Unterstützung *vor* dem Dolmetscheinsatz verwendet werden, beispielsweise zur Glossarerstellung bzw. zum Terminologiemanagement, ebenso *während* des Einsatzes, dank einer Transkripterstellung in Echtzeit mittels automatischer Spracherkennung (ASR – ‚automatic speech recognition') bzw. Speech-to-Text-Technologie, sowie auch *nach* dem Einsatz zur Nachbereitung bzw. Terminologiepflege (vgl. Fantinuoli, 2018a). Das Forschungsinteresse galt dabei bisher in erster Linie der Entwicklung und dem Einsatz von CAI-Tools beim Simultandolmetschen, oftmals unter Verwendung von automatischer Spracherkennung sowie mit einem Schwerpunkt auf der Verdolmetschung von Zahlen als Problemauslöser (vgl. dazu die Studien von Cheung & Li, 2022; Defrancq & Fantinuoli, 2020; Fantinuoli, 2016, 2017; Pisani & Fantinuoli, 2021; Rodríguez González et al., 2023).

Die technologischen Fortschritte ermöglichen nun auch den Einsatz von Tools zum computergestützten Konsekutivdolmetschen (CACI – ‚computer-assisted consecutive interpreting'), welches aufgrund seiner Neuheit noch großen Forschungsbedarf aufweist (Chen, 2017; Chen & Kruger, 2022). Die vorliegende Masterarbeit beschäftigt sich daher mit einer erst vor kurzem entwickelten und im Rahmen einer Studie getesteten Online-Anwendung namens Sight-Terp, welche das computergestützte Konsekutivdolmetschen ermöglichen und erleichtern soll (Ünlü, 2023). Ünlü analysierte im durchgeführten Experiment die Verdolmetschungen von 12 Studierenden im konsekutiven Arbeitsmodus mit und ohne Sight-Terp im Sprachenpaar Englisch-Türkisch und befragte im Anschluss daran die Teilnehmenden per Fragebogen nach ihren Eindrücken (2023: 76ff.). Dabei fand Ünlü (2023: 95) heraus, dass die

Genauigkeit der Verdolmetschungen unter Verwendung von Sight-Terp anstieg, dass die Anwendung des Tools jedoch auf den Aspekt der Flüssigkeit der Wiedergabe einen negativen Einfluss ausübte.

Die vorliegende Masterarbeit versteht sich in ihrem empirischen Teil als eine Replikation der Masterarbeit von Ünlü (2023). Eine Replikation ist die Wiederholung eines Experiments unter sehr ähnlichen Bedingungen, um zu überprüfen, ob die daraus gewonnenen Erkenntnisse dieselben oder andere sind (Gile, 2016: 221), wobei in der vorliegenden Arbeit die Anwendung von Sight-Terp beim Konsekutivdolmetschen in der Sprachkombination Englisch-Deutsch sowohl quantitativ als auch qualitativ eingehender erforscht werden soll. Die gewonnenen Erkenntnisse sollen einen weiteren Beitrag zur Erforschung des computergestützten Konsekutivdolmetschens liefern. Die erste Forschungsfrage lautet:

1) Inwiefern kann die Verwendung von Sight-Terp beim Konsekutivdolmetschen im Sprachenpaar Englisch-Deutsch im Vergleich zur ausschließlichen Verwendung der klassischen Notizentechnik zu einer Verbesserung der Qualität führen, gemessen anhand der Genauigkeit der Wiedergabe von Sinneinheiten sowie der Verdolmetschung ausgewählter Problemauslöser?

Basierend auf der Definition von CAI-Tools nach Fantinuoli (2018a: 155f.), nämlich als Computersoftware, die der Verbesserung der Qualität und Produktivität dient, sowie anhand des aktuellen Forschungsstandes (vgl. Abschnitt 3.3.3) sei zudem die folgende Hypothese aufgestellt: Die Anwendung Sight-Terp verbessert die Qualität beim konsekutiven Dolmetschen im Sprachenpaar Englisch-Deutsch, gemessen anhand der Anzahl der fehlerfreien Sinneinheiten, sowie der Anzahl der korrekt gedolmetschten Problemauslöser im Vergleich zur Verdolmetschung ohne Sight-Terp.

Die zweite Forschungsfrage lautet:

2) Wie arbeiten Dolmetschstudierende insbesondere nach eigener Aussage mit Sight-Terp und wie schätzen sie ihre eigene konsekutive Dolmetschleistung mit bzw. ohne Sight-Terp ein?
    a. Inwiefern ändert sich die Dauer der Verdolmetschung mit Sight-Terp?
    b. Wie fertigen die Teilnehmenden ihre Notizen bei der gleichzeitigen Verwendung von Sight-Terp an?

Um mögliche Antworten auf diese Forschungsfragen zu erhalten, werden im empirischen Teil der vorliegenden Arbeit im Sinne des Mixed-Methods-Ansatzes, der aufgrund der Kombination aus den erhobenen quantitativen und qualitativen Daten eine breitere Perspektive auf die Forschungsthematik zulässt (vgl. Hale & Napier, 2014: 210f.), einerseits ein Experiment, andererseits ein Leitfaden-Interview (vgl. Helfferich, 2014: 560) mit fortgeschrittenen Dolmetschstudierenden ab dem 4. Semester des Zentrums für Translationswissenschaft durchgeführt. Das Experiment umfasst ein sich wiederholendes Messungsverfahren (‚repeated measures design'), wobei die gleichen Teilnehmenden beiden gemessenen Bedingungen abwechselnd ausgesetzt werden (vgl. Mellinger & Hanson, 2016: 7).

Dieses Verfahren bedeutet in der vorliegenden Arbeit zunächst die konsekutive Verdolmetschung einer ersten Rede mit Sight-Terp und ggf. Notizen, sowie einer zweiten Rede ausschließlich mit der klassischen Notizentechnik für TN1 sowie in umgekehrter Reihenfolge für TN2, usw., jeweils aus dem Englischen ins Deutsche, wobei dafür ausgewählte englische Originalreden aus der Masterarbeit von Ünlü (2023) herangezogen werden. Im Anschluss daran werden die Teilnehmenden zu ihren Eindrücken bei der Anwendung von Sight-Terp befragt, um neben quantitativen auch qualitative Daten zu erheben.

Die gewonnenen quantitativen Daten werden schließlich ausgewertet, indem in den Ausgangsreden die Sinneinheiten, also bedeutungstragende Einheiten, in der Definition nach Lederer (2014: 18) mittels einer erweiterten Fehleranalyse nach Zhao (2021: 94ff.) erkannt werden. Diese werden dann in den konsekutiven Verdolmetschungen mit bzw. ohne Sight-Terp anhand einer Kategorisierung in Auslassungen, Hinzufügungen, Ersetzungen und eines stufenweisen Erkennens des zugehörigen Schweregrads auf ihre Genauigkeit hin bewertet. Die Problemauslöser (‚problem triggers'), wie z. B. Eigennamen, Zahlen, Aufzählungen, Termini, Abkürzungen, etc. in der Definition nach Gile (1995: 172ff.; 2009: 171), werden ebenso erkannt und dahingehend beurteilt, ob sie von den Teilnehmenden verdolmetscht (1) oder nicht verdolmetscht (0) wurden. Letztlich werden die qualitativen Daten, die im Rahmen des Leitfaden-Interviews gewonnen werden konnten, mithilfe der qualitativen Inhaltsanalyse nach Mayring (2022) Kategorien laut einem eigens zu diesem Zwecke angefertigten Kodierleitfaden zugeteilt und ausgewertet.

Im theoretischen Teil der vorliegenden Arbeit sollen die Grundlagen anhand eines Überblicks über die verschiedenen (kombinierten) Arbeitsmodi (vgl. Pöchhacker, 2016) beim Dolmetschen präsentiert werden, wie z. B. simultan, konsekutiv, Vom-Blatt-Übersetzen, Flüsterdolmetschen, Dolmetschen mit Text im simultanen und konsekutiven Modus, Simkons mit dem Smartpen, etc. Für die zumeist vorherrschenden Arbeitsmodi simultan und konsekutiv

kommen ebenso kognitive Aspekte anhand des (aktualisierten) Effort-Modells nach Gile (1985/ 2020) bzw. des Cognitive-Load-Modells nach Seeber (2011) zur Sprache. Diese werden aufgrund ihres Ursprungs zunächst für den simultanen Modus skizziert, um dann einen Übergang zum konsekutiven Modus herzustellen. Im Rahmen des Abschnitts über den konsekutiven Arbeitsmodus wird auch die Notationstechnik im klassischen Sinne, also mit Stift und Notizblock (vgl. Pöchhacker 2016: 192f.), anhand der grundlegenden Vertreter sowie eine technologiegestützte Variante, wie diese beim Digital Notepad zu finden ist (vgl. Goldsmith, 2017, 2018), vorgestellt.

Anschließend werden Analyseperspektiven für Verdolmetschungen anhand möglicher Zugänge zur Qualität erläutert, welche wiederum auf Qualitätskriterien und Bewertungsmethoden aufgeteilt sind. Letztere werden mittels Problemauslösern und Sinneinheiten sowie anhand des Aspekts der Genauigkeit im Zuge der Fehleranalyse dargestellt.

Das darauffolgende Kapitel behandelt die bisherigen computergestützten Anwendungen und technologische Hintergründe zu den dabei angewandten Funktionalitäten künstlicher Intelligenz sowie den diesbezüglichen Forschungsstand, erneut aufgeteilt in den simultanen und konsekutiven Arbeitsmodus. Im Zuge der Darlegung des aktuellen Forschungsstands beim computergestützten Konsekutivdolmetschen wird insbesondere die Studie von Ünlü (2023) unter Verwendung der Anwendung Sight-Terp erläutert, welche im daran anknüpfenden empirischen Teil im Zuge der Replikation herangezogen wird.

1. Grundlagen

Dieses Kapitel umreißt einige der am häufigsten auftretenden Dolmetschmodi in lautsprachlicher Form, wobei diese im Sinne der Arbeitsmodi beim Dolmetschen nach Pöchhacker (2016: 18) verstanden werden und keinen Anspruch auf Vollständigkeit erheben, sondern einem ersten Überblick dienen sollen.

Die Tätigkeit des Dolmetschens lässt sich zunächst als eine einmalige Wiedergabe einer einmalig präsentierten ausgangssprachlichen Nachricht in einer Zielsprache definieren (Pöchhacker, 2016: 11), weshalb als wichtiges Abgrenzungsmerkmal zum Übersetzen beim Dolmetschen kaum die Möglichkeit besteht, das Translat im Anschluss an seine Produktion zu verändern.

Die Arbeitsmodi werden grob gesagt dahingehend eingeteilt, wann in Bezug auf die Äußerung der Ausgangsrede und mit welchen Mitteln gedolmetscht wird. Eine grundlegende Unterscheidung wird dabei zumeist zwischen dem konsekutiven und dem simultanen Dolmetschen getroffen. Konsekutivdolmetschen umfasst die Wiedergabe einer ausgangssprachlichen Botschaft in einer Zielsprache, nachdem diese in ihren Segmenten jeweils vollständig geäußert wurde, wohingegen das Simultandolmetschen während des Vortrags der ausgangssprachlichen Botschaft erfolgt (Pöchhacker, 2016: 18).

Ein weiterer Unterschied ist, dass Konsekutivdolmetscher*innen den Text als Ganzes analysieren und die Bedeutung dahinter verstehen können (Seleskovitch, 1978: 123), wohingegen Simultandolmetscher*innen aufgrund des (nahezu) gleichzeitigen Verdolmetschens auf viel mehr eigenes Wissen zurückgreifen müssen, um den Sinn der gesamten Aussage zu verstehen (Seleskovitch, 1978: 125) bzw. Fehlendes zu ergänzen.

Der konsekutive Modus wird meist noch eingeteilt in den klassischen konsekutiven Modus (‚long consecutive'), bei dem auch die Notizentechnik (vgl. Abschnitt 1.2.2) angewandt wird, im Unterschied zum kurzen Konsekutivdolmetschen (‚short consecutive'), welches Satz für Satz und daher ohne Notizentechnik möglich ist (Dam, 2010: 75).

Das Simultandolmetschen ist zwar einerseits sehr eng mit den technologischen Fortschritten verbunden, da es nicht zuletzt aus der Notwendigkeit der Kosten- und Zeitersparnis im Rahmen mehrsprachiger Veranstaltungen, wie etwa für Konferenzen, entwickelt wurde (Russo, 2010: 333), impliziert aber nicht zwangsläufig die Anwendung technischer Geräte. Das Flüsterdolmetschen wird beispielsweise simultan ausgeübt, wobei die dolmetschende Person durch leises Sprechen nahe der Zuhörer*innen die ausgangssprachlichen Textelemente in der Zielsprache wiedergibt.

Insbesondere die Erweiterung der Reichweite ist ein Argument dafür, weshalb bei größerem Dolmetschbedarf ein Technologieeinsatz erforderlich wird. So kann Flüsterdolmetschen auch technologiegestützt umgesetzt werden, nämlich mittels Flüsteranlage, die zumeist ein Mikrofon und Kopfhörer für die Zuhörer*innen umfasst (Pöchhacker, 2016: 19f.).
Das Konsekutivdolmetschen bleibt jedoch oft die günstigste und flexibelste Form des Dolmetschens und findet daher nach wie vor bei unterschiedlichen Veranstaltungen Anwendung (Dam, 2010: 76).

Daneben gibt es auch Mischformen zwischen den Arbeitsmodi. Hier ist die Kombination aus akustischem und textuellem Input anzuführen, wie beispielsweise das Vom-Blatt-Übersetzen. Dabei wird eine schriftliche Botschaft in einer Ausgangssprache in einer mündlichen, oder nicht-schriftlichen, Form in der Zielsprache ausgedrückt (Čeňková, 2010: 320). Dies ist z. B. notwendig, wenn schriftliche Dokumente, wie mitgebrachte Berichte oder Urkunden, komplett oder teilweise neben dem eigentlichen Vortrag präsentiert werden sollen.

Eine weitere hybride Form ist das Simultandolmetschen mit Text, bei dem die Dolmetscher*innen das vorzutragende Redemanuskript für die Verdolmetschung zur Verfügung gestellt bekommen. Da diese Form eine visuelle und eine akustische ausgangssprachliche Komponente enthält und die akustisch wiedergegebenen die tatsächlich zu verdolmetschenden Segmente darstellen, können sich Abweichungen vom schriftlichen Text als Problemauslöser herausstellen (vgl. Abschnitt 2.2.1); daher wird dieser Arbeitsmodus als besonders komplex erachtet (Pöchhacker, 2016: 20; Seeber, 2015: 87).

Die technologischen Entwicklungen im Rahmen des computergestützten Dolmetschens (vgl. Kapitel 3) ermöglichen das Entstehen weiterer Formen, wie etwa SimConsec und SightConsec (Corpas Pastor, 2021b: 96), also konsekutives Dolmetschen in Kombination mit dem simultanen Modus (vgl. dazu auch ‚Smartpen', Abschnitt 3.3.1) bzw. in Kombination mit dem Vom-Blatt-Übersetzen schriftlichen Materials. Auf diese technologischen Mischformen wird in den Abschnitten zu Kapitel 3 näher eingegangen.

Im Folgenden werden zunächst die kognitiven Aspekte im Arbeitsmodus simultan und konsekutiv erläutert, um zu verstehen, welche Prozesse beim (computergestützten) Dolmetschen zum Tragen kommen.

1.1 Simultandolmetschen

1.1.1 Kognitive Prozessmodelle

Simultandolmetschen stellt eine komplexe kognitive Fähigkeit dar, weil neben der Gleichzeitigkeit der Darbietung einer ausgangssprachlichen Botschaft in der Zielsprache auch weitere Aspekte wie die Koordination des gleichzeitigen Zuhörens und Sprechens bzw. auch gutes Allgemeinwissen zum Tragen kommen (Russo, 2010: 333).

An dieser Stelle seien zwei Modelle zur Beschreibung der kognitiven Belastung zunächst für das Simultandolmetschen angeführt, nämlich insbesondere das Effort-Modell nach Gile (1985) und das darauf aufbauende kognitive Belastungsmodell (‚Cognitive Load Model') nach Seeber (2011).

In seiner Grundform beschreibt das Effort-Modell die Komponenten des Zuhörens und der Analyse (‚L' für ‚listening and analysis'), der Produktion (‚P' für ‚production') und des Gedächtnisses (‚M' für ‚memory') als quasi-mathematische Gleichung (Gile, 1985: 44f.). Gile (2009: 168) beschreibt auch eine Koordinationsanstrengung (‚C' für ‚coordination') für die Handhabung aller Anstrengungen beim Simultandolmetschen mit- bzw. untereinander. Die Gleichung ist eine Addition der Efforts und lautet:

$$SI = L + P + M + C$$

Die Addition verdeutlicht, dass beim Simultandolmetschen alle diese Anstrengungen zeitlich beinahe gleichzeitig bzw. möglichst zusammen erfolgen, sodass im Umkehrschluss beim stärkeren Belasten eines Efforts die anderen zwangsläufig weniger Aufmerksamkeit bekommen, da die Gleichung letzten Endes auf den gleichen Wert, also in diesem Fall die maximal erreichbare kognitive Belastung, kommen muss.

Das Effort-Modell von Gile hat es sich ebenso zum Ziel gesetzt, das Konfliktpotenzial in Form von Problemauslösern (‚problem triggers', vgl. Abschnitt 2.2.1) zu erklären. Gile (1995: 172ff.; 2009: 171f.) merkt an, dass es Fehlersequenzen geben kann, bei denen ein Problemauslöser ein späteres unproblematisches Segment negativ beeinflusst. Jedoch kann eine verminderte Qualität schwer mit dem Problem assoziiert werden, welches diese verursacht, wenn das fehleranfällige, aber eigentlich unproblematische, Segment eventuell von dem verursachenden Segment weit entfernt liegt (Gile, 1995: 175; 2009: 172).

Seeber (2011) nimmt in seinem kognitiven Belastungsmodell in Anlehnung an Giles Effort-Modell hingegen auch auf ein mögliches Konfliktpotenzial im Zusammenspiel der Efforts untereinander und auf mögliche Interferenzen bei einer Überschneidung der Anstrengungen Bezug. Außerdem versucht Seeber (2011) die kognitive Belastung zu quantifizieren, indem jeder Anstrengung bzw. den Überschneidungen der jeweiligen Efforts Konfliktkoeffizienten zugewiesen werden (Seeber, 2011: 188f.). Seeber (2011: 186) argumentiert in diesem Sinne auch, dass die Kombination von kognitiven Aufgaben beim Simultandolmetschen Probleme auslösen kann, die bei ihrer getrennten Ausführung nicht entstehen würden.

1.1.2 Kognition und computergestütztes Dolmetschen

Gile (2020) führt aufgrund der technologischen Fortschritte eine weitere Komponente für sein Modell ein. Im Zusammenspiel zwischen „Mensch und Maschine“ (‚human-machine interaction‘, kurz HMI) seien beim Simultandolmetschen neben den oben beschriebenen Efforts auch die zusätzliche Anstrengung der Interaktion zwischen Mensch und Maschine bzw. HMI sowie die Aufnahme (‚R‘ für ‚reception‘) sowohl in akustischer als auch in visueller Form in das Modell zu integrieren (Gile, 2020: 11). Dies sieht folgendermaßen aus:

$$SI = L + P + M + C + HMI + R$$

Hier wird sichtbar, dass das computergestützte Dolmetschen nach Fantinuoli (2018a: 168) per Definition multimodal ist, da zu den üblichen Stimuli, wie oben erwähnt, noch weitere hinzukommen, wie beispielsweise die Suche nach der richtigen Information auf dem Bildschirm. An dieser Stelle sei angemerkt, dass weitere Anstrengungen mit den anderen wiederum in einem Zusammenspiel stehen und ggf. neu verteilt werden müssen. Fantinuoli (2017: 3) erkennt, dass etwa die Suche nach einem Ausdruck während des computergestützten Simultandolmetschens zu einer kognitiven Überlastung führen kann.

Prandi (2018: 56) schätzt ebenso die kognitive Belastung beim Simultandolmetschen mit CAI-Tools im Vergleich zur traditionellen Methode als höher ein, meint aber, dass Tools wie InterpretBank die kognitive Belastung im Vergleich zu „first-generation tools“ wie Word und Excel vermindern (vgl. dazu Guo et al. 2023: 93; Fantinuoli 2018a: 164).

In Hinblick auf das computergestützte Dolmetschen erwähnt Fantinuoli (2016: 48) weiter, dass die kognitive Belastung, die durch den Einsatz eines Tools zu der bereits vorhandenen Anstrengung des Dolmetschens an sich hinzukommt, so gering wie möglich sein sollte. Das kann erzielt werden, indem etwa Benutzerfreundlichkeit in den Vordergrund der

Entwicklung rückt, welche insbesondere durch eine klare Funktionsweise beim Arbeiten mit der Anwendung gegeben ist.

Eine weitere Abhilfe zur beschriebenen Problematik kann der Einsatz von künstlicher Intelligenz bieten, da mittels automatischer Spracherkennung und maschineller Übersetzung eine wichtige Vorarbeit zur kognitiven Entlastung geleistet werden kann (vgl. Kapitel 3). Das Vorhandensein technologischer Unterstützung könnte trotz einer (anfänglich) gestiegenen kognitiven Last beim Umgang mit dem Tool somit andererseits die in der oben beschriebenen Gleichung angesprochene maximale Belastungskapazität erhöhen und positive Auswirkungen auf (einzelne) Efforts haben.

Fantinuoli (2017: 4) macht ebenfalls in der Anwendung von automatischer Spracherkennung das Potenzial aus, die zusätzlich erforderliche kognitive Anstrengung der HMI zu reduzieren und erwähnt u. a. die automatische Transkription von Problemauslösern wie Zahlen, Abkürzungen und Eigennamen. Diese Überlegung ist auch für das im Folgenden dargestellte Konsekutivdolmetschen von Relevanz.

1.2 Konsekutivdolmetschen

1.2.1 Kognitive Prozessmodelle

Das Effort-Modell wurde zwar ursprünglich für das Simultandolmetschen entwickelt, jedoch wurde ein solches auch für den konsekutiven Modus konzipiert (Gile, 1995: 178ff.; 2009: 175ff.) und sei an dieser Stelle angeführt. Gile (2009: 175) unterteilt konsekutives Dolmetschen in zwei Phasen, nämlich jene des Verstehens, also des Zuhörens und des Notierens (‚L' für ‚listening' und ‚N' für ‚note-taking'), und jene der Spracherzeugung oder Reformulierung, wobei die Komponenten des Kurzzeitgedächtnisses (‚M') und der Koordination (‚C') darin ebenso enthalten sind. Dies führt zu folgender Gleichung:

$$\text{Interpreting} = L + N + M + C$$

Gile (2009: 175f.) erkennt, dass das Zuhören und die Analyse (hier ‚L') jenem gleichen Effort des Simultandolmetschens entsprechen, ebenso wie eine Ähnlichkeit der jeweiligen Gedächtnisanstrengung besteht, die sich beim Konsekutivdolmetschen auf die Zeit zwischen dem Hören und Notieren bezieht. Die Produktionsphase bezieht sich nun zunächst auf das Notieren und erst in einer zweiten Phase werden die zieltextsprachlichen Segmente produziert.

Die zweite Gleichung lautet:

$$\text{Interpreting} = \text{Rem} + \text{Read} + \text{P} + \text{C}$$

Dabei versteht Gile (2009: 176) unter ‘Rem‘ (kurz für ‚remembering‘) das Wiederaufrufen der in der ersten Phase erhaltenen Informationen im Langzeitgedächtnis und ist daher von dem bis dato beschriebenen Kurzzeitgedächtnis ‚M‘ zu unterscheiden. Zusätzlich müssen neben der Produktions- bzw. Reformulierungs- (‚P‘) und Koordinationsanstrengung (‚C‘) auch noch die gemachten Notizen gelesen werden (‚read‘). Gile (2009: 176) erwähnt, dass Notizen, die einfachen Regeln folgen (vgl. Abschnitt 1.2.2), dem Gedächtnis auf die Sprünge helfen können sowie dass das Tempo des Vortrags in der zweiten Phase im Unterschied zur ersten bzw. zum Simultandolmetschen nicht mehr von der vortragenden Person bestimmt wird, sondern beim konsekutiven Dolmetschen selbst gewählt werden kann. Gile (2009: 176) schreibt, dass die Verarbeitungskapazität daher lediglich in der ersten Phase des Zuhörens und Notierens zu einer Sättigung aufgrund von Zeitdruck führen kann.

1.2.2 Kognition und computergestütztes Dolmetschen

Da sich die vorliegende Arbeit mit der computergestützten konsekutiven Verdolmetschung beschäftigt, kann die Gleichung zur kognitiven Leistung unter Verwendung von CAI-Anwendungen um die beim Simultandolmetschen erwähnte Anstrengung ‚HMI‘ (Gile, 2020: 11), nämlich in diesem Fall betreffend die Interaktion zwischen der dolmetschenden Person und der Softwareanwendung Sight-Terp, sowie um die Aufnahme ‚R‘ zur folgenden Gleichung erweitert werden:

$$\text{Interpreting} = \text{Rem} + \text{Read} + \text{P} + \text{C} + \text{HMI} + \text{R}$$

Die Antwort auf die Frage, inwiefern die zusätzlichen Anstrengungen des konsekutiven Dolmetschens mit maschineller Integration zu einer Überlastung und somit zu einer negativen Beeinträchtigung der Verdolmetschung führen könnten, wird folglich davon abhängen, wie die gemachten Notizen bzw. das automatische Transkript und die maschinelle Übersetzung einer Rede durch das CAI-Programm angefertigt werden und wie bzw. ob Dolmetscher*innen diese in der zweiten Phase in Anspruch nehmen.

Chen & Kruger (2022: 413) erkennen beispielsweise, dass ein zur Verfügung stehender maschinell übersetzter Text die kognitive Belastung verringert, was einen positiven Effekt

auf die flüssige Wiedergabe der Verdolmetschung in der Produktionsphase haben kann (vgl. Forschungsstand, Abschnitt 3.3.3). Die Beschreibung der kognitiven Belastung im neuen Arbeitsmodus CACI (‚computer-assisted consecutive interpreting') wurde vom beschriebenen Effort-Modell inspiriert und umfasst ebenfalls zwei Phasen (vgl. Abbildung 1). Die erste Phase beinhaltet nach Chen & Kruger (2022: 405) das Zuhören (‚listening' und ‚analysis'), die Gedächtnisleistung (‚memory') und das laute Nachsprechen (‚respeaking') in ein Spracherkennungssystem, welches somit von der Rede ein automatisches Transkript anfertigt und dieses in ein maschinelles Übersetzungssystem einspeist. In der zweiten Phase erfolgt dann die konsekutive Verdolmetschung (‚production') unter Zuhilfenahme des ausgangssprachlichen Transkripts und der zielsprachigen Übersetzung anhand des Lesens (‚reading') des mittels Spracherkennung angefertigten Transkripts bzw. der davon angefertigten maschinellen Übersetzung sowie des Erinnerns (‚remembering') an das Vorgetragene. In beiden Phasen kommt zudem die Koordination (‚coordination') der Efforts zum Tragen.

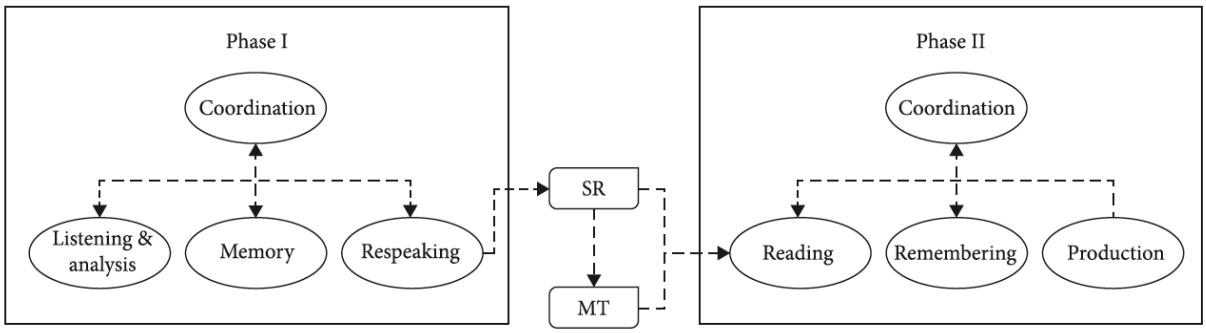


Abbildung 1: Arbeitsmodus CACI (‚computer-assisted consecutive interpreting') (Chen & Kruger, 2022: 405)

Die in Abbildung 1 dargestellte Vorgehensweise im Arbeitsmodus des computergestützten Konsekutivdolmetschens macht Unterschiede zum Konsekutivdolmetschen ohne Technologieeinsatz ersichtlich. Insbesondere beziehen sich die Efforts der Phase zwei nunmehr auf die dank der automatischen Spracherkennung und maschineller Übersetzung angefertigten Mitschriften und es stellt sich die Frage nach dem Einsatz der klassischen Notizentechnik.

Der folgende Abschnitt beschäftigt sich daher nicht zuletzt aufgrund dieser neuer, kombinierter Arbeitsmodi genauer mit der Entwicklung und dem Einsatz der Notizentechnik beim konsekutiven Dolmetschen, um herauszufinden, inwiefern Notizen mit oder ohne Technologieeinsatz die Dolmetscher*innen (zusätzlich) ent- bzw. belasten können.

1.2.3 Notizentechnik

Im folgenden Abschnitt wird zunächst auf die klassische sowie auch auf die von modernen Technologien beeinflusste Notizentechnik Bezug genommen, da sie einen untrennbaren Bestandteil im konsekutiven Arbeitsmodus darstellt.

Die klassische oder traditionelle Art des Konsekutivdolmetschens wird dabei in diesem Sinne als jene Methode verstanden, die einen Notizblock und einen Stift als hauptsächliche Hilfsmittel vorsieht (Pöchhacker 2016: 192f.).

In seinem Grundlagenwerk definiert Herbert (1968: 32) das Ziel der Notizentechnik als eine effektive Ergänzung des Gedächtnisses, wobei die dolmetschende Person die Notizen in Folge einer logischen Analyse der Rede anfertigen sollte (1968: 34). In Hinblick auf die Sprache der Notizen erwähnt Herbert (1968: 36f.), dass diese jene sein soll, in der die Verdolmetschung erfolgen wird, jedoch spricht er neben dieser sprachgebundenen Notation auch von dem Vorteil sprachunabhängiger Symbole bei der Notizenanfertigung (1968: 37).

Rozan (1984: 13) beschreibt später in seinem Grundlagenwerk die Notizentechnik in sieben Prinzipien und liefert zahlreiche praktische Anwendungsbeispiele. Darin werden Regeln zum Notieren von Abkürzung, Bindewörtern, Verneinungen, Betonungen sowie erstmals eine vertikale und verschobene Notizentechnik präsentiert. Letztlich soll der Gedanke hinter dem Gesagten und nicht das Wort als sprachlicher Ausdruck dieses Gedankens beim Konsekutivdolmetschen notiert werden. Rozan (1984: 15) stellt außerdem fest: „À chaque instant de la prise de notes il faut se concentrer sur l'idée maîtresse et la transposer de façon simple et directe (…)".

Zudem führt Rozan (1984: 27) auch die Verwendung von Symbolen ein, jedoch sollten diese keine Aneinanderreihung darstellen, die es erneut zu entziffern gilt. Vielmehr sollen die Symbole möglichst schnell dem Gedächtnis helfen, sich an den Ausgangstext zu erinnern. Einige dieser Symbole umfassen unter vielen anderen Pfeile für Anstiege und Rückgänge, mathematische Symbole wie (Un-)Gleichheitszeichen für die (Un-)Gleichstellung zweier Gedanken sowie Unterstreichungen zum Hervorheben von besonders wichtigem Inhalt.

Die beiden Techniken nach Herbert (1968) und Rozan (1984) sind zunächst eher präskriptiv orientiert, da sie eine korrekte Notationsweise vorgeben, die es zu verwenden gilt. Matyssek (2012: 37f.) postuliert hingegen eine „sprachlose Notizentechnik", das bedeutet eine „Abkehr vom Wort als Worthülse" und präsentiert in seinem Werk Notationssymbole, die von (werdenden) Dolmetscher*innen möglichst sprachungebunden und somit universell für das Konsekutivdolmetschen erlernt und angewandt werden können. Matyssek (2012: 40)

beschreibt die Schaffung einer „geistigen Basis“ als Aufgabe der Notizentechnik und präsentiert dementsprechend möglichst vielfältig anwendbare Symbole im Kontext zahlreicher Beispiele sowie ein Nachschlageregister der meistgebrauchten Symbole zum Erlernen bzw. als Inspiration für eigene Symbole. Damit wendet sich diese Notationstechnik der deskriptiven Methode zu, da die vorgeschlagenen Techniken in erster Linie Empfehlungen darstellen.

Gillies (2017) greift viele der vorgestellten zugrundeliegenden Aspekte auf. Der Sinn bzw. die Idee einer Äußerung steht beim Notieren im Vordergrund, ebenso wie eine Aufteilung in die Grundstruktur Subjekt-Verb-Objekt (2017: 37ff.). Außerdem wird eine diagonale Notationsweise bevorzugt (2017: 43ff.) und das Kapitel über Konnektoren (2017: 60ff.) macht deutlich, dass die durch einen Trennstrich entstehende linke Spalte für eben diese Bindewörter und weitere für die Verdolmetschung wichtige Elemente verwendet werden kann.

Die Gemeinsamkeit der angesprochenen Werke besteht nicht zuletzt darin, dass ein Zugang zur Notizentechnik postuliert wird, der das Notieren eines Gedankens bzw. einer Sinneinheit anstatt eines Wortes hervorhebt. Diese Überlegung greift die „théorie du sens“ auf, in deren Rahmen Seleskovitch (1968: 161f.) behauptet, dass der Gedanke, nicht das Wort, beim Dolmetschen ausgedrückt werden soll. Eine dahingehend angepasste Notation erleichtert somit die Produktionsphase. Daneben geben die praktischen Anwendungsbeispiele in den präsentierten Werken zu erkennen, dass die bereits erwähnte vertikale bzw. leicht versetzte Notation die Verdolmetschung erleichtert. Zudem empfehlen die meisten oben angeführten Werke eine vertikale Trennlinie, die zwei Spalten entstehen lässt und der Aufteilung von etwa Personalpronomen, Verben, Modalverben, etc. dienen soll und somit einen besseren Überblick schafft. Zur Frage, in welcher Sprache notiert werden sollte, liefern die Ergebnisse einer Studie von Abuín Gonzàlez (2012: 66) aufschlussreiche Erkenntnisse, da herausgefunden wurde, dass professionelle Dolmetscher*innen im Vergleich zu Studierenden öfter in der Zielsprache als in der Ausgangssprache notierten.

An dieser Stelle sei nun kritisch angemerkt, dass das im Zuge des empirischen Teils angewandte Tool Sight-Terp vom ausgangssprachlichen Text ein automatisches Transkript als Fließtext erzeugt, das für die konsekutive Verdolmetschung herangezogen werden kann (vgl. Ünlü, 2023: 65). Hier stellt sich die Frage, ob die (alleinige) Verwendung eines Fließtexts in Anbetracht der „théorie du sens“ bzw. der oben beschriebenen Überlegungen zur Notation in Symbolen, etc. für das konsekutive Dolmetschen geeignet ist. Dieser Problematik ist sich auch Ünlü (2023: 70) bewusst und er integrierte daher in Sight-Terp eine automatische Nummerierung und Markierungen, um eine gewisse Strukturierung der Ausgangsrede als Anhaltspunkt im Sinne der Notizentechnik zu bewahren.

Nun wurde auch die Notizentechnik vom technologischen Fortschritt beeinflusst. So kommt es beim computergestützten Konsekutivdolmetschen vor, dass statt Stift und Papier ein Tablet und ggf. ein dafür vorgesehener Eingabestift als elektronische Hilfsmittel zum Notieren verwendet werden und die Notizentechnik technologisch eingesetzt werden kann.

Auch das in dieser Arbeit angewandte Tool Sight-Terp bietet die Möglichkeit eines digitalen Notizblocks (‚digital notepad', vgl. Ünlü, 2023: 73ff.), welches jedoch im empirischen Teil im Gegensatz zum klassischen Notizblock und Stift nicht zum Einsatz kommt.

Goldsmith (2017: 41ff.) führte in diesem Zusammenhang eine Studie durch, bei der die Nutzer*innen von Tablets danach gefragt wurden, welche Funktionalitäten Tablets, elektronische Anwendungen und Stifte für die Erstellung der Notizen haben sollten, und erstellte eine Liste mit wünschenswerten Merkmalen. In den Antworten geht insbesondere die Verlässlichkeit der elektronischen Anwendungen als ein entscheidendes Kriterium hervor.

Die Erkenntnis im Zusammenhang mit der Forderung nach Verlässlichkeit ist auch für das in der vorliegenden Arbeit angewandte Tool Sight-Terp von Interesse, da technische Probleme im Zweifelsfall gegen das computergestützte Konsekutivdolmetschen allein und für die technikfreie und dadurch unabhängigere klassische Notizentechnik mit Papier und Stift oder zumindest für eine Kombination beider Arbeitsweisen sprechen (vgl. dazu Abschnitt 5.2).

An dieser Stelle sei die Studie von Goldsmith (2018) angeführt, in deren Rahmen eine qualitative Untersuchung zur Verwendung des Tablets unter professionellen Dolmetscher*innen durchgeführt wurde. Dabei fand Goldsmith (2018: 356) heraus, dass Tablets zwar bei Dolmetscheinsätzen bereits auf vielfältige Art und Weise Anwendung finden, andererseits aber auch, dass Aspekte wie die Benutzerfreundlichkeit, die Zuverlässigkeit und die Reaktion der Kund*innen dazu führen, dass in vielen Fällen statt zu dem elektronischen Tablet doch lieber zu Stift und Papier gegriffen wird.

Abschließend sei auch die Studie von Chen (2017) angeführt, bei der die Notizentechnik beim Konsekutivdolmetschen zwischen Chinesisch und Englisch aus einer kognitiven Perspektive beleuchtet wurde, indem ein digitaler Stift zum Einsatz kam. Dabei wurde herausgefunden, dass eine gute Verdolmetschung sowohl Informationen in Form von gemachten Notizen als auch jene im Gedächtnis abruft (Chen, 2017: 17) und dass Symbole und Abkürzungen die kognitiven Anstrengungen bzw. Efforts reduzieren (2017: 19f.), was wiederum notwendig ist, um die begrenzte Verfügbarkeit dieser kognitiven Ressourcen effektiv aufzuteilen, damit die Notizen eine Unterstützung und keine zusätzliche Belastung für die Dolmetscher*innen darstellen.

Dies führt schlussendlich wieder auf die oben beschriebene Problematik zurück, inwiefern Transkripte und im übergeordneten Sinne Anwendungen für das computergestützte Konsekutivdolmetschen eine kognitive Belastung für Dolmetscher*innen darstellen. Sind sie eine Ergänzung angefertigter Notizen, ebenso wie die Notizen eine Ergänzung des Gedächtnisses sein sollten und erlernt werden müssen (vgl. Gile, 1995: 182), so tragen sie womöglich zu einer Entlastung bei. Wird der Umgang mit diesen Tools jedoch nicht erlernt und erprobt, so könnten sie zu einer kognitiven Überanstrengung führen, entweder wenn Anwender*innen nicht damit umzugehen wissen und die Tools so eine Ablenkung darstellen bzw. auch, wenn sich Dolmetscher*innen zu sehr auf die technische Unterstützung verlassen und diese plötzlich wegfällt, was die kognitive Last kurzzeitig vervielfachen kann (vgl. dazu Abschnitt 3.2.5 und die Ergebnisse von Defrancq & Fantinuoli, 2020: 98).

Die kognitive Belastung steht in einer Wechselbeziehung mit der Dolmetschleistung bzw. -qualität, die mit erhöhter Anstrengung tendenziell abnimmt (vgl. Gieshoff & Albl-Mikasa, 2022: 210). Um die Qualität unter diversen Aspekten bewerten zu können werden im Folgenden mögliche Qualitätskriterien und Bewertungsmethoden vorgestellt.

2. Qualität

Die wissenschaftliche Auseinandersetzung mit der Qualitätsbeurteilung begann grundsätzlich mit dem Aufkommen des Simultandolmetschens und ist immer noch stärker im simultanen, als im konsekutiven Arbeitsmodus vorzufinden (Kalina, 2002: 122). Die Analyse einer Verdolmetschung kann dabei unterschiedliche Sichtweisen umfassen. Im Folgenden werden die prozess-, produkt- und personenorientierten Perspektiven umrissen.

Zunächst spricht Toury (1995: 13) von einer prozess- und produktorientierten Forschung in der Translationswissenschaft. Die voraussichtliche Position und Funktion eines Translats bestimmt seine textuell-linguistische Umsetzung als Produkt, was wiederum den Prozess zur Schaffung dieses Produkts sowie seine Beziehung zum Ausgangstext leitet. Die Auffassung von Nutzer*innen davon, was Qualitätsmerkmale einer Verdolmetschung sind, betrifft dabei das Produkt und der Vergleich zwischen Original, Verdolmetschung und den Umständen betrifft den Prozess (Collados Aís & García Becerra, 2015: 373).

Fantinuoli (2018a: 167) bezieht sich in der Forderung nach produkt- und prozessorientierter Forschung im Bereich des computergestützten Dolmetschens auch auf diese Einteilung und versteht unter produktorientierter Forschung in diesem Zusammenhang quantitative und qualitative Analysen gedolmetschter Texte, die zumeist komparativ, also mittels eines Vergleichs des Ausgangs- und Zieltexts, etwa dank kontrastiver Diskursanalysen bzw. des Vergleichs des (Nicht-)Vorhandenseins eines entsprechenden Tools, erfolgen. Letzterer Ansatz, der den Vergleich einer Verdolmetschung mit und ohne Softwareanwendung untersucht, wurde auch im Rahmen dieser Arbeit gewählt. Fantinuoli (2018a: 168) spricht hierbei auch von der Messung des Einflusses von CAI-Tools auf die Leistung der Dolmetscher*innen und somit bezieht sich diese Perspektive auf das Endprodukt der Verdolmetschung.

Andererseits erwähnt Fantinuoli (2018a: 168), dass die prozessorientierte Forschung kognitive Aspekte umfasst, beispielsweise als Antwort auf die Frage, wie Konferenzdolmetscher*innen mit multimodalen sensorischen Informationen umgehen, wenn sie Informationen während des Dolmetschens auf einem Bildschirm suchen und lesen müssen. Diese Perspektive beschäftigt sich also mit der Phase während des Dolmetschens.

Eine weitere Perspektive umfasst die Bewertung der dolmetschenden Person. In der von Pradas Macías (2023: 82) durchgeführten Befragung wurden 13 Aspekte zur Bewertung von Dolmetscher*innen abgefragt, darunter etwa inwiefern diese selbstbewusst, kompetent oder auch sichtbar waren. Daneben wurden auch produktorientierte, also die

Verdolmetschung betreffende, Fragen gestellt, etwa zur Geschwindigkeit, zum Sprachregister und zum Verständnis, zur Kohärenz und ebenso zur Treue zum Original.

Kalina (2015: 17ff.) zeigt in einem Überblick über Zugänge zur Qualitätsbestimmung, dass deren Grundlage ursprünglich vor allem der Vergleich und das Identifizieren von Abweichungen zwischen der transkribierten Ausgangsrede, also dem Original, und der Verdolmetschung bildete. Das Fehlen authentischer Materialien stellte lange Zeit eine der größten Herausforderungen für diese Art von Qualitätsbestimmung dar.

In einem weiteren Verlauf rückte der Zieltext und dessen Analyse vermehrt in den Vordergrund. Shlesinger (1997: 128) empfiehlt die Analyse des Zieltexts auf drei Ebenen: intertextuell als Vergleich zwischen Ausgangs- und Zieltext, intratextuell basierend auf der Verdolmetschung selbst und instrumental je nach ihrer Nützlichkeit. Kalina (2015: 19f.) schlägt vor, nicht nur den Zieltext bzw. die Verdolmetschung, sondern auch den Ausgangstext, d. h. den zugrundeliegenden Redebeitrag, auf individueller Basis zu berücksichtigen.

Eine weitere Analysemethode findet sich bei Lamberger-Felber (2001: 39ff.), die in inhaltsorientierte, prozessorientierte und formorientierte Forschungsebenen unterteilt. Die inhaltsorientierte Forschung umfasst die Überprüfung von Verdolmetschungen mit dem jeweiligen Ausgangstext auf ihre Genauigkeit hin (vgl. Fehleranalyse, Abschnitt 2.2.3), die prozessorientierte Forschung bezieht sich auf die Analyse der Handlungen während des Dolmetschens und die formorientierte Forschung beschäftigt sich u. a. mit der Geschwindigkeit, aber auch mit der Kohäsion und vermehrt sprachspezifischen Aspekten.

Trotz der Tatsache, dass es keine einheitliche Methode zur Beschreibung des Produkts der Verdolmetschung in einer bestimmten Situation gibt (Pöchhacker, 1994a: 235), bildet die produkt- bzw. inhaltsorientierte Analyse die Basis für die vorliegende Arbeit. In einem ersten Schritt werden daher mögliche Kriterien zur Analyse bzw. Bestimmung von Qualitätsaspekten angesprochen. Im Anschluss daran werden Methoden zur Segmentierung von Redebeiträgen und den davon angefertigten Verdolmetschungen vorgestellt. Diese umfassen das Erkennen von Problemauslösern und Sinneinheiten, wobei die Forderung nach einer dem Original treuen bzw. einer genauen und vollständigen Verdolmetschung die am weitesten verbreitete Art der Bewertung darstellt und somit berücksichtigt wird (vgl. Pöchhacker, 2016: 140f.). Schließlich wird für die Analyse der Genauigkeit die Fehleranalyse vorgestellt.

2.1 Bestimmung von Qualitätskriterien

Obwohl im Zuge der Analyse von Verdolmetschungen zumeist Aspekte der Qualität zur Sprache kommen, gibt es keine einheitlichen Kriterien bzw. keine Einigung darüber, wie diese zu definieren sind (Collados Aís & García Becerra, 2015: 368). Die Qualitätsbeurteilung einer Verdolmetschung stellt nicht zuletzt wegen einer Vielzahl an möglichen Parametern ein schwieriges Unterfangen dar (Kalina, 2002: 125). Kalina (2015: 18) führt das Fehlen gemeinsamer Qualitätskriterien auf die besonderen Umstände zurück, in der die Verdolmetschung durchgeführt wird. Die Schwierigkeit der Qualitätsbewertung ist nämlich nicht zuletzt dadurch gegeben, dass das Endprodukt, also die Verdolmetschung, in seiner Natur flüchtig ist und nicht erneut beurteilt werden kann, geschweige denn als ein Produkt festgelegt werden kann (Garzone, 2002: 107; Pöchhacker, 1994a: 235). Daher stellt sich eine Beurteilung auf produktorientierter Ebene meist als schwer heraus. Zudem wäre eine schriftliche Analyse der Verdolmetschung laut Garzone (2002: 107) eine methodisch falsche Vorgehensweise, da Faktoren wie die Prosodie bzw. die Präsentation ebenso von starker Bedeutung sind und in der Verschriftlichung unberücksichtigt bleiben. Für eine Beurteilung ist jedenfalls vorab eine Definition jener Beurteilungskriterien, anhand derer die Dolmetschleistung gemessen werden kann, von entscheidender Bedeutung.

Die Ergebnisse einer Befragung von Bühler (1986) unter Mitgliedern des internationalen Verbands der Konferenzdolmetscher*innen (AIIC) zeigen, welche Kriterien diese für die Bewertung der Qualität heranziehen bzw. worauf sie besonders achten. Dabei zeigt sich, dass beinahe alle inhalts- und sprachbezogenen Kriterien, wie die Flüssigkeit, der logische Zusammenhang der Aussage (‚logical cohesion of utterance'), die Sinnübereinstimmung (‚sense consistency with the original message'), die Vollständigkeit und die Verwendung der korrekten Grammatik und Terminologie als sehr wichtig eingestuft werden und dass Kriterien wie der Akzent oder eine angenehme Stimme als wünschenswert, aber nicht als grundlegend bewertet werden (Bühler, 1986: 233). In Bezug auf die Qualitätsbeurteilung durch Endverbraucher*innen merkt Bühler (1986: 233) kritisch an, dass eine Unterscheidung zwischen den Bedürfnissen und den Erwartungen letzterer getroffen werden muss, da die Erwartungen bei oberflächlichen Kriterien, etwa betreffend eine angenehme Stimme, oft unrealistisch ausfallen würden.

Eine von Pöchhacker & Zwischenberger (2010: 2ff.) durchgeführte ähnliche Befragung als Replikation der Studie von Bühler (1986) unter AIIC-Mitgliedern zeigt, welchen ausgewählten Qualitätskriterien diese bei einer simultanen Verdolmetschung Wichtigkeit

beimessen. Dabei zeigen die Ergebnisse zu diesem späteren Zeitpunkt, dass neben den von Bühler (1986) erwähnten Kriterien der Sinnübereinstimmung mit dem Ausgangstext (‚sense consistency with the original') sowie des logischen Zusammenhangs (‚logical cohesion') vor allem formbezogene Kriterien, wie eine korrekte Terminologie und auch eine flüssige Darbietung als wichtigste Merkmale qualitativer Verdolmetschungen betrachtet werden.

Auch García Becerra (2015: 130) untersucht die Evaluierung simultaner Verdolmetschungen anhand der Aspekte der sogenannten „allgemeinen" Qualität sowie formbezogener Aspekte, inhaltsbezogener Aspekte, der Treue und des Verständnisses.

Kalina (1994: 229) bezieht sich wiederum bei der Untersuchung von Verdolmetschungen von Studierenden und professionellen Dolmetscher*innen auf weitere Kriterien, welche auch Fehler und Aspekte der Vollständigkeit und Relevanz betreffen.

Gile (2009: 171f.) unterscheidet bei der Verminderung der Qualität aufgrund von Problemen mit der Verarbeitungskapazität (vgl. Problemauslöser 2.2.1) einerseits zwischen einer Abweichung des Inhalts in der Zielsprache, gemessen an Fehlern, Auslassungen etc., sowie andererseits jener der Darbietung, die den sogenannten sprachlichen Output bzw. auch Aspekte wie die Stimme und die Intonation betreffen.

Im Zusammenhang mit der Bewertung von Studierenden führt Riccardi (2002: 118ff.) Kriterien auf der Makro- und auf der Mikroebene an. Erstere umfasst Merkmale wie Äquivalenz (vgl. Abschnitt 2.2), Genauigkeit, Adäquatheit und letztere betrifft etwa prosodische und morphosyntaktische Merkmale. Außerdem wird ein Evaluierungsbogen für Konferenzdolmetschprüfungen präsentiert (vgl. Riccardi, 2002: 124ff.).

Setton & Motta (2007: 202) machen eine qualitative Verdolmetschung anhand dreier Sichtweisen aus. Erstens, je nach dem impliziten Konsens unter professionellen Dolmetscher*innen darüber, was eine gute Verdolmetschung ist. Zweitens, je nach den Informationen über die Erwartungen und Reaktionen der Nutzer*innen der Verdolmetschung und drittens, je nach der Messung einer Äquivalenz zwischen der Ausgangs- und der Zielsprache, auf semantischer Ebene oder auf Wirkungsebene. Als hauptsächliche Bewertungsmethode würden dabei zumeist die ‚senior peers', also professionelle Dolmetscher*innen mit langjähriger Erfahrung, herangezogen, insbesondere was den Konsens in Bezug auf eine qualitative Verdolmetschung betrifft.

Kalina (2002: 125) stellt mögliche qualitätsbeeinflussende Faktoren übersichtlich und modusübergreifend dar. Dabei werden die Merkmale zur Qualitätsbestimmung in Kategorien eingeteilt und so zwischen den Aspekten des semantischen Inhalts, der sprachlichen Leistung und der Darbietung unterschieden. Kriterien des semantischen Inhalts umfassen u. a.

Konsistenz, einen logischen Zusammenhang, Vollständigkeit und Genauigkeit. Jene der sprachlichen Leistung etwa die grammatikalische und terminologische Korrektheit und jene der Darbietung beispielsweise die Stimmqualität und das Artikulieren. Dabei betont Kalina (2002: 125), dass die Faktoren ineinander verflochten sind und von sich ändernden Bedingungen abhängen.

Collados Aís et al. (2007: 215) unterteilen die am häufigsten herangezogenen Kriterien zur Qualitätsbestimmung in vier Blöcke mit absteigender Priorität. Diese umfassen zunächst die Sinnübereinstimmung und die Kohäsion, danach die Vollständigkeit, Terminologie und Flüssigkeit, anschließend die Diktion, den Stil und die Grammatik sowie letztlich die Aussprache, die Stimme und den Akzent.

Moser-Mercer (1996: 44) postuliert eine bestmögliche Qualität (‚optimum quality'), welche eine komplette und genaue Wiedergabe des Originals umfasst, ohne dieses zu verzerren, sowie auch mögliche außersprachliche Informationen berücksichtigt, die wiederum von äußeren Bedingungen abhängen. Diese betreffen etwa die physische Umgebung, die Themenkomplexität sowie insbesondere Merkmale des Redebeitrags und der Präsentation.

Die Bewertung wird nicht zuletzt von einer beurteilenden Person bestimmt. Je nachdem, wer die Verdolmetschung beurteilen soll, können die herangezogenen Qualitätskriterien dementsprechend anders ausfallen (vgl. Kalina, 2002: 121). Bühler (1986: 231) definiert die beurteilende Person dabei als Lehrende einer akademischen Ausbildungsstätte, als professionelle Organisation, etwa im Zuge von Akkreditierungen, als Arbeitgeber*innen, wie z. B. die Organisator*innen einer Konferenz oder als Endverbraucher*innen, also die Zuhörer*innen der Verdolmetschung.

Da die Beurteilung einer Form von Genauigkeit, Sinnübereinstimmung bzw. Vollständigkeit oder umgekehrt der Abweichung von Verdolmetschungen von dem entsprechenden Ausgangstext jedoch häufig mit der Qualität in Verbindung gesetzt wird (vgl. Pöchhacker, 2016: 142), beziehen sich die hier vorgestellten Methoden auf diesen Aspekt.

2.2 Methoden zur Qualitätsbewertung

In Hinblick auf die Qualität der Verdolmetschung führt Seleskovitch (1968: 196) das Kriterium der Genauigkeit beim Konsekutivdolmetschen insofern an, als in diesem Arbeitsmodus die Zuhörer*innen beide Teile, also Ausgangs- und Zieltext, separat hören und ggf. nach ihrer Genauigkeit beurteilen können (vgl. dazu auch Gile, 1983: 242). Beim simultanen Modus könnten hingegen nur professionelle Dolmetscher*innen erkennen, ob die Punkte des Ausgangstexts auch im Zieltext, also in der Verdolmetschung, zu finden sind bzw. diesen

entsprechen (Seleskovitch, 1968: 198). Seleskovitch (1968: 166) fordert in jedem Fall eine sogenannte „absolute Treue“ (‚fidélité absolue‘).

Gile (1983: 241) bezeichnet die übergreifende Qualitätsbewertung der Verdolmetschung u. a. als ein Werkzeug zur Bestimmung der relativen Bedeutung, die die beurteilende Person der Informationstreue (‚fidélité informationnelle‘) zuschreibt.

Die angesprochene Forderung nach Treue lässt sich mit der Überlegung in Verbindung bringen, dass die Qualität der Leistung nicht zuletzt im Zuge eines Vergleichs des Ausgangs- und Zieltexts eruiert werden kann (vgl. Setton & Motta, 2007: 202). Bühler (1986: 233) schreibt hierzu: „The criteria of *sense consistency with the original message* (5) and *completeness of interpretation* (6), which are essential for interlingual communication and hence also the quality of interpretation, can only be judged by comparison with the original.” Ein intertextueller, also textübergreifender, Vergleich der Verdolmetschung mit dem Ausgangstext ist daneben auch für das Kriterium der „logical cohesion“ der zu bevorzugende Ansatz (Pöchhacker, 1994a: 241).

An dieser Stelle ist auch die sogenannte Äquivalenz zu erwähnen, deren Evaluierung sich primär in der Übersetzungstheorie herausgebildet hat (vgl. Lederer, 2014: 61ff.). Lederer (2014: 45) unterscheidet: „(…) equivalence exists between texts, correspondences between linguistic elements (…)“ und hebt hervor, dass das Ziel einer Übersetzung eine übergreifende Äquivalenz sein soll und nicht allein die Übereinstimmung auf der Ebene linguistischer Elemente, wie Wörtern. Diese Überlegung der Äquivalenz zwischen der Ausgangs- und Zielsprache ist auch bei der Evaluierung von Verdolmetschungen aufgrund von Genauigkeit bzw. Abweichungen und Fehlern ein implizierter Ansatz (vgl. Pöchhacker, 2016: 59).

Zur Bewertung einer Äquivalenz bzw. auch ‚sense consistency‘ sei die Unterscheidung von Setton & Motta (2007: 203) angeführt, die diese auf drei Ebenen unterteilen, nämlich zunächst als Vergleich der Verdolmetschung mit einem sogenannten ‚gold standard‘, also einer erstrebenswerten idealen bzw. im Vorhinein als solches zu definierende Verdolmetschung, mittels Fehleranalyse (‚error analysis‘), wofür Barik (1975) den Grundstein legte, oder durch die Segmentierung der Ausgangs- und Zieltexte in semantische Informationseinheiten (vgl. ‚propositional analysis‘). Bei letzterer Methode werden die Ausgangs- und Zieltexte in ihre kleinsten sinntragenden Informationseinheiten unterteilt, wobei eine semantische Informationseinheit im Sinne der Propositional Analysis nach Ding (2017: 19) als „smallest unit that can express a complete meaning, which can be in the form of a word, a phrase, a clause or a sentence” definiert werden kann.

In der durchgeführten Studie unterteilt Ding (2017: 24) die semantischen Informationseinheiten in Form von Propositionen in Prädikate bzw. Verben, Modifikationen bzw. Modalverben, und Bindewörtern bzw. Konjunktionen und listet diese im Anhang (2017: 34ff.) detailliert auf. Folglich versucht eine derartige Analysemethode aufgrund ihrer Granularität den semantischen Informationsgehalt im Ausgangstext kleinstmöglich zu segmentieren und die Verdolmetschung auf dessen Vorhandensein zu analysieren. Als hauptsächlicher Kritikpunkt an diesem Analyseverfahren kann festgehalten werden, dass diese Art der Segmentierung sprachgebunden ist und die Problematik der semantischen Vergleichbarkeit zwischen zwei Sprachen nicht lösen kann (Pöchhacker 2016: 142).

In der experimentellen Forschung wird die Genauigkeit als Aspekt einer qualitativen Verdolmetschung dabei oft als eine von anderen Parametern abhängige Variable eingesetzt, wobei die Einflussparameter, also die unabhängigen Variablen, neben weiteren Aspekten der Arbeitsmodus, die Arbeitsbedingungen und die Sprachrichtung darstellen können (Pöchhacker, 2016: 142).

Die Frage nach einem äquivalenten Informationsgehalt sowie dessen Definition in zwei unterschiedlichen Sprachen ist dementsprechend schwer zu beantworten. Das bedeutet grundsätzlich auch, dass beispielsweise eine Segmentierung in der Ausgangssprache Englisch und eine Segmentierung des gedolmetschten deutschen zielsprachlichen Textes nicht auf der gleichen Ebene erfolgen kann, da sich die semantischen Informationseinheiten nicht zuletzt aufgrund des Satzbaus bzw. der Syntax entsprechend verschieben und so auch sichtbar wird, dass ein Informationsgehalt im Zieltext zusammengezogen werden kann, der dem Ausgangstext in seiner Bedeutung entspricht.

Eine Kritik an der beschriebenen Analyse der Genauigkeit mittels Segmentierung ebenso wie einen möglichen Lösungsansatz liefern Gieshoff & Albl-Mikasa (2022: 212f.), da eine solche Methode bislang nicht die Entwicklung der Sinnübereinstimmung bzw. die allgemein Kohäsion berücksichtigt hätte. Sie präsentieren daher eine Methode, die die Genauigkeit ebenfalls anhand von Einheiten misst, welche auf der Propositional Analysis aufbauen, wobei sie jedes Element je nach seiner Funktion im Ausgangstext einer Kategorie zuordnen, insbesondere um die Kohäsion und die metasprachliche Ebene zu berücksichtigen. Die Weiterentwicklung erlaubt den Einsatz statistischer Analyseverfahren, eine Vielzahl an Bewertungen zu unterschiedlichen Zeitpunkten und berücksichtigt die mündliche Natur einer Verdolmetschung. Das bedeutet, dass etwa Auslassungen von Wiederholungen, eine geänderte Reihenfolge und eine Zusammenziehung mehrerer Elemente nicht als Fehler gewertet werden.

Die Mündlichkeit sowie die sogenannte allgemeine Sinnübereinstimmung (vgl. dazu Bühler, 1986: 233) wird in Anlehnung an Gieshoff & Albl-Mikasa (2022: 213) ebenso im vorliegenden empirischen Teil berücksichtigt, da Abweichungen in der Verdolmetschung, wenn diese nicht zu einer Sinnverzerrung führten, nicht als Fehler, wie sie im Zuge der Fehleranalyse (vgl. Abschnitt 2.2.3) beschrieben werden, gewertet werden.

Im Gegensatz zur kleinstmöglichen Einheit, wie diese im Zuge der Propositional Analysis postuliert wird, kann eine Segmentierung zur Bestimmung der Genauigkeit schließlich auch anders erfolgen. Die von Bühler (1986: 233) angesprochene „sense consistency" und die „equivalence of sense", bei der die Absicht des Vortragenden über die dolmetschende Person zu den Zuhörenden gelangen soll (vgl. Pöchhacker, 2016: 92), führt wiederum zu folgender Überlegung von Seleskovitch (1968: 161) zurück: „Ce que dit l'interprète est donc, en principe, indépendant de la langue étrangère". Lederer (2014: 18) erkennt, dass eine Sinneinheit zur Schaffung von Äquivalenz führen kann, wobei eine kleinere Einheit nur eine Entsprechung auf Wortebene zulassen würde.

Eine weitere Segmentierung bzw. Beurteilungsgrundlage umfasst jene nach Problemauslösern und ist u. a. in der Studie von Mankauskiené (2018a: 6ff.) zu finden, bei der neben einer Klassifizierung der Problemauslöser eine deskriptiv-analytische sowie komparative Methode eingesetzt wird, um die Auswirkungen erkannter Problemauslöser auf die Verdolmetschung zu erforschen.

Im Folgenden werden daher zwei Methoden zur Segmentierung der Ausgangsreden und der zugehörigen Verdolmetschung vorgestellt. Zunächst werden die Problemauslöser erläutert und im Anschluss daran wird eine Einteilung in Sinneinheiten nach Lederer (2014: 18) getroffen, wobei die Fehleranalyse als komparative Methode den Abschluss bildet.

2.2.1 Problemauslöser

Eine erste Auseinandersetzung mit Herausforderungen in Form von Problemen bzw. Schwierigkeiten kommt von Nord (1991), wobei hier zwischen Problemen im Sinne objektiver Schwierigkeiten einer Ausgangsrede, die für alle Dolmetscher*innen gleichermaßen eine potenzielle Herausforderung darstellen können, und Schwierigkeiten, die subjektiv und daher individuell für die jeweilige dolmetschende Person sind, unterschieden wird. In der vorliegenden Arbeit wird diese detaillierte Unterteilung nicht getroffen.

Problemauslöser (‚problem triggers', kurz PT) nach Gile (1995: 172; 2009: 171) erfordern den Einsatz einer erhöhten Verarbeitungskapazität (vgl. Efforts) durch die dolmetschende Person. Gile (1995: 171f.) führt darunter und in Bezug auf das Simultandolmetschen

ein schnelles Vortragstempo, eine hohe Dichte, externe Faktoren, wie umgebenden Lärm, unbekannte Wörter sowie nicht zuletzt Zahlen und Eigennamen an. Daneben können sich Problemauslöser u. a. als Abkürzungen, Aufzählungen und starke Akzente manifestieren (Gile, 2009: 171). Problemauslöser erhöhen somit die benötigte Verarbeitungskapazität beim Dolmetschen, was zu einer kognitiven Überlastung, Aufmerksamkeitsproblemen und letztlich zu Leistungseinbußen führen kann. Gile (1995: 181f.) stellt ebenso fest, dass die Problemauslöser in der ersten Phase des Konsekutivdolmetschens dieselben sind wie jene beim Simultandolmetschen, wohingegen der Unterschied in der Notationstechnik liegt, die die im Folgenden aufzubringende Gedächtnisleistung weiter beeinflusst (vgl. Abschnitt 1.2.1).

In diesem Zusammenhang sind auch die sogenannten Named Entities (kurz NE) zu nennen. Darunter werden jene Elemente in einem Satz verstanden, die Namen von Personen, Organisationen bzw. Orte sind. Als Teilaspekt der angesprochenen Problemauslöser, wie etwa Eigennamen, stellt das Erkennen der Named Entities eine wichtige Aufgabe von Systemen zur Informationsgewinnung dar (vgl. Tjong Kim Sang & De Meulder, 2003: 142).

Mankauskiené (2018a: 17) klassifiziert Problemauslöser weiter je nach ihrer Herkunft bzw. Ursache. Dabei wird zwischen senderbezogenen Problemauslösern, wie dem Redetempo und dem Akzent der vortragenden Person, den textbezogenen Problemauslösern, die auf die Lexik, die Syntax bzw. den Satzbau und die Semiologie des Vorgetragenen zurückzuführen sind, sowie jenen Problemauslösern, die sich auf die dolmetschende Person mit ihrem jeweiligen Hintergrundwissen beziehen, und letztlich technischen Problemauslösern unterschieden.

Auch Seeber (2015) erwähnt die Problemauslöser im Zusammenhang mit dem simultanen Arbeitsmodus. Seeber (2015: 87) trifft eine Unterscheidung in Geschwindigkeit und Dichte als eine Kategorie, sowie jeweils Zahlen, Syntax und Akzente, die nicht zuletzt aufgrund des Phänomens Englisch als Lingua Franca (‚ELF') eine Herausforderung für Dolmetscher*innen mit Englisch als Arbeitssprache darstellen.

Zur optimalen Redegeschwindigkeit empfiehlt der internationale Verband der Konferenzdolmetscher*innen AIIC (2015) eine Vortragsgeschwindigkeit für zu verdolmetschende Reden von „3 Minuten pro Seite à 40 Zeilen", das macht bei einer Standardzeilenlänge von 12 Wörtern eine maximale Geschwindigkeit von bis zu 160 Wörtern pro Minute. Beim Konsekutivdolmetschen im Rahmen von Konferenzen beträgt die durchschnittliche Vortragsgeschwindigkeit nach Seleskovitch (1978: 31) 150 Wörter pro Minute, wobei dieser Wert als sehr viel Konzentration erfordernd und anstrengend eingeschätzt wird.

Neben den Wörtern pro Minute schreibt Seeber (2015: 85) in diesem Zusammenhang, dass langsam vorgetragene, aber sehr dichte Reden als schnell wahrgenommen werden

können. Dichte Reden könnten dabei durch ein enges Aneinanderreihen von Problemauslösern erkannt werden, die aufgrund ihrer kognitiven Belastung die Rede schneller erscheinen lassen. In der vorliegenden Arbeit werden die Problemauslöser schließlich nach Gile (1995: 171f.) definiert und für den konsekutiven Arbeitsmodus herangezogen.

Zahlen stellen besondere Problemauslöser dar, da sie vom üblichen verbalen Input abweichen. Sie haben nämlich, bis auf sehr bekannte, etwa historische Datumsangaben, keine konzeptuelle Darstellung, sie können also nicht im Vorhinein erkannt, das heißt antizipiert, werden (Gile, 1995: 174; Seeber, 2015: 85f.). Zudem sind sie auch nicht anderweitig bildlich darstellbar. Ein erschwerender Faktor ist auch, dass Zahlen aufgrund ihrer Kürze schnell ausgesprochen werden und dass die Wahrscheinlichkeit hilfreicher Redundanzen, also Wiederholungen ähnlicher Informationen, entsprechend gering ist. Das alles erhöht die benötigte kognitive Verarbeitungskapazität. Aus diesem Grund wurde speziell im Zuge des computergestützten Simultandolmetschens viel im Kontext der Verdolmetschung von Zahlen als Problemauslöser geforscht, um zu erkennen, inwiefern die Leistung beim technologiegestützten Dolmetschen verbessert werden kann (vgl. Abschnitt 3.2.5).

Betreffend die Syntax erwähnt Seeber (2015: 86), dass Sprachen, in denen das Verb am Ende eines Satzes stehen kann, wie das beispielsweise im Deutschen der Fall ist, nicht den Verstehensprozess an sich beeinträchtigen. Jedoch führen syntaktische Unterschiede zwischen Sprachen zu einer erhöhten Inanspruchnahme von Verarbeitungskapazität und dies nicht zuletzt beim Simultandolmetschen, bei dem die Sätze vor der Produktion der Verdolmetschung selten zu Ende gehört werden können (vgl. Abschnitt 1.1.1).

Da im konsekutiven Arbeitsmodus die Sätze vor der Produktion der Verdolmetschung gehört werden und sich die Analyse hierbei auf die Verdolmetschung und nicht etwa auf den Prozess des Notierens während des Zuhörens der Rede konzentriert (vgl. dazu Chen, 2017), wird die Syntax als möglicher Problemauslöser in dieser Arbeit nicht in die Analyse mitaufgenommen. Zur Feststellung eines erhöhten Bedarfs an Verarbeitungskapazität kann jedoch neben der Syntax auch die Lesbarkeit beurteilt werden, beispielsweise anhand des Gunning Fog Index (vgl. Abschnitt 4.2.1). Dieser Index stellt eine Berechnungsmethode dar, die aufgrund der Länge der Wörter die Schwierigkeit eines Textes beurteilt. Diese Analyse ist insofern als Mittel anzusehen, um die beiden Ausgangsreden auf eine vergleichbare Vortragsweise zu überprüfen, als schwer lesbare Texte aufgrund komplex zusammengesetzter Wörter potenziell die Vortragsweise beeinträchtigen könnten. Die Berechnung des Index erfolgt in der vorliegenden Arbeit daher unabhängig davon, dass die syntaktischen Unterschiede im Sprachenpaar Englisch-Deutsch nicht weiter untersucht werden.

Die Überlegung von Gile (1995: 174f.; 2009: 171f.) in Bezug auf Fehlersequenzen, die dann entstehen, wenn Problemauslöser kognitiv derart anspruchsvoll sind, dass auch nachfolgende unproblematische Segmente negativ beeinflusst werden können, ist auch an dieser Stelle von Relevanz, da die erwähnten Problemauslöser aufgrund einer erhöhten kognitiven Belastung die Wiedergabe weiterer Sinneinheiten beeinträchtigen könnten. Gile (2009: 172) erkennt jedoch, dass der Zusammenhang zwischen einem problematischen und dem verursachenden Segment schwer feststellbar ist.

Aus diesem Grund beschränken sich die in den Ausgangsreden analysierten und als Basis für die Bewertung herangezogenen Problemauslöser auf die oben erwähnten Eigennamen, Abkürzungen, Zahlen, Aufzählungen sowie auch die Termini in Form von selten vorkommendem und daher vorzubereitendem Wortschatz und die Dichte bzw. das Vortragetempo (vgl. Abschnitt 4.2.3).

Neben dem Erkennen von Problemauslösern werden im nächsten Abschnitt auch die sogenannten Sinneinheiten präsentiert, die ebenso bei der darauffolgenden Fehleranalyse von Relevanz sein werden.

2.2.2 Sinneinheiten

Die Überlegung von Seleskovitch (1968: 160ff.), dass die dolmetschende Person zwischen den Prozessen des Hörens und des Ausdrückens eines zu vermittelnden Gedankens diesen zunächst im Zuge einer Deverbalisierung verstehen muss, dass also der Gedanke vor seiner Formulierung nonverbal sei und nicht zuletzt dank unterschiedlichen Vorwissens erneut ausgedrückt werden kann, führt zur sogenannten „théorie du sens“. Die erwähnte Metaebene des Sinns ist nicht zuletzt für die sinnbasierte Arbeitsweise im konsekutiven Arbeitsmodus von entscheidender Bedeutung (vgl. Pöchhacker, 2015: 64f.).

Lederer (2014: 18) bezeichnet konsekutives Dolmetschen als Demonstration des kognitiven Gedächtnisses und Simultandolmetschen als schrittweisen Sinnaufbau, wobei Wörter in gewissen Zeitabschnitten Verstehen auslösen. Jeder Zeitabschnitt schafft eine sogenannte geistige Einheit (‚mental unit‘), die von den Zuhörenden erkannt werden muss. Lederer (2014: 18) definiert Sinneinheiten als Produkt der Verschmelzung von Wortbedeutungen und kognitiven Ergänzungen.

Lederer (2014: 18f.) spricht beim schrittweisen Aufbau einer Sinneinheit (‚unit of sense‘, kurz SE) beim Simultandolmetschen davon, dass diese zunächst bewusst wahrnehmbar ist, bevor sie durch die Bildung weiterer Sinneinheiten in den Hintergrund tritt. Zudem hätten Sinneinheiten keine vorgegebene Länge und jene Zuhörer*innen bzw.

Dolmetscher*innen, die aufgrund ihres individuellen Hintergrundwissens mit einem gegebenen Thema vertraut sind, müssten nicht auf das Ende des akustischen Inputs warten, um zu verstehen, worum es in einer Aussage geht. Lederer behauptet (2014: 18) ferner: „Thus, depending on the knowledge which addressees bring to bear, a speech may be redundant for some or too elliptical for others". Dementsprechend haben Sinneinheiten keine fixen Grenzen und werden von den jeweiligen Zuhörenden in unregelmäßige Intervalle unterteilt (Seleskovitch & Lederer, 1989: 247).

Eine Sinn- bzw. Bedeutungseinheit (‚unit of sense') manifestiert sich nun im Rahmen einer Rede in Form kleinster, deverbalisierter Elemente, deren Erkennen eine Entsprechung von Segmenten eines Textes ermöglicht (vgl. Lederer, 2014: 18f.) und somit für eine gedankliche Struktur sorgt. Diese Struktur gilt es nicht zuletzt im konsekutiven Arbeitsmodus zu erfassen, um sie wiederum für das Anfertigen der Notizen und die anschließende konsekutive Verdolmetschung nützen zu können (vgl. Abschnitt 1.2.2).

Ein Beispiel für die Unterteilung in Sinneinheiten als Analyse- bzw. Bewertungsmethode ist in der Studie von Orlando (2014: 44) zu finden (vgl. Abschnitt 3.3.1), in der die Messung der Genauigkeit anhand der Anzahl der vollständig gedolmetschten Sinneinheiten, welche Fakten und Ideen repräsentieren, erfolgt.

Ünlü (2023: 89) unterteilt im Rahmen der Datenanalyse ebenso die für das Experiment herangezogenen englischen Ausgangstexte in Sinneinheiten. Auf Basis dieser Strukturierung sowie unter Verwendung von Methoden der Propositional Analysis erfolgt somit eine Beurteilung der Genauigkeit.

Eine ähnliche Herangehensweise, bei der Sinneinheiten anhand einer kleinstmöglichen Segmentierung im Sinne der Propositional Analysis zum Erkennen der strukturellen Bedeutung (vgl. ‚structural meaning') hervorgehoben werden, ist daneben u. a. bei Dillinger (1994: 163) zu finden (zur Kritik an der Methode vgl. Pöchhacker, 2016: 142).

Die Unterteilung der Ausgangsrede in Sinneinheiten ist neben dem simultanen gerade im konsekutiven Arbeitsmodus naheliegend, da die Notizentechnik (vgl. Abschnitt 1.2.2) auf dem Notieren von Konzepten und einer gedanklichen Struktur hinter dem Gesagten basiert und Sinneinheiten die Grundlage für verwendete Abkürzungen und Symbole bilden können. Gillies (2017: 37f.) empfiehlt für das Notieren von Ideen, sich die Antwort auf die Frage „who does what to whom" zu überlegen, wobei hier die sogenannten W-Fragen, also wer, was, wem, etc., klar zum Vorschein treten. Diese Methode wird auch in der vorliegenden Arbeit angewandt (vgl. Abschnitt 4.2.3).

Das Hauptaugenmerk der Beurteilung liegt schlussendlich auf einem Vergleich der Entsprechung der erkannten Segmente der Ausgangs- und Zieltexte. Genauigkeit wird aufgrund einer Segmentierungsmethode, die Sinneinheiten, aber auch Problemauslöser (vgl. Abschnitt 2.1.1) erkennt, sowie durch die Ermittlung jenes Inhalts der Verdolmetschung gemessen, welcher vom Geäußerten abweicht. Diese Überlegung leitet zur Genauigkeit und zur Messung letzterer anhand von Methoden der Fehleranalyse über.

2.2.3 Fehleranalyse

Die im methodischen Teil dieser Arbeit zur Anwendung kommende Bewertung der Verdolmetschungen orientiert sich an einer vergleichenden Analyse zwischen dem in Sinneinheiten bzw. Problemauslöser segmentierten Originaltranskript und der Verdolmetschung. Die Genauigkeit wird dabei aufgrund einer erweiterten Fehleranalyse nach Zhao (2021), welche insbesondere auf Barik (1975) aufbaut, ermittelt.

Barik leistete dabei einen ersten Beitrag zur Bewertung von Verdolmetschungen, wurde jedoch für diese einerseits zu ausgeklügelte Methode kritisiert (Pöchhacker: 2016: 142f.), bei der andererseits weitere grundlegende Aspekte einer Verdolmetschung, wie etwa die Prosodie, unbeachtet bleiben (Kalina, 2015: 17). Da diese Methode jedoch die Basis der Fehleranalyse bildet, wird sie im Folgenden erläutert. Die Methode nach Barik (1975: 275ff.) unterteilt Fehler in drei Arten von Abweichungen (‚translation departures').

Erstens, Auslassungen, also Elemente, die zwar in der Ausgangsrede, nicht jedoch in der Verdolmetschung vorhanden sind. Irrelevante Wiederholungen oder Füllwörter sind hier ausgenommen. Weiters werden Auslassungen unterteilt in einzelne, weniger kritische Auslassungen, jene, die auf einen Verständnisfehler hinweisen und zu einem Bedeutungsverlust führen, sowie Auslassungen aufgrund eines verspäteten Einsatzes der dolmetschenden Person oder auch jene, die auf eine Zusammenziehung mehrerer Elemente zurückzuführen sind und die somit leicht von ihrer ursprünglichen Bedeutung abweichen.

Zweitens, Hinzufügungen, also Elemente, die in der Verdolmetschung auftauchen, nicht aber im Ausgangstext. Dabei werden weniger sinntragende Hinzufügungen, wie Adverbien, von ausführlicheren unterschieden, ebenso wie Hinzufügungen, die eine nicht ausgedrückte Beziehung zwischen Sätzen implizieren sowie auch solche, die einen Satz ausgeschmückt beenden, ohne Wesentliches dazu beizutragen.

Drittens, Ersetzungen und Fehler, bei denen die dolmetschende Person etwas ausdrückt, das (so) nicht im Ausgangstext vorkam. Diese werden in leichte semantische Fehler und schwere semantische Fehler aufgegliedert. Die erste der beiden Fehlerkategorien enthält

zwar die Kernaussage, ist aber nicht ganz korrekt. Die zweite Kategorie ändert wesentlich die Bedeutung des Gesagten, entweder aufgrund von Homonymen, also ähnlich klingenden Wörtern in zwei Sprachen, die aber eine andere Bedeutung tragen, oder aber aufgrund von einer falschen Bezugnahme (‚confusion of reference') oder da es sich um einen offensichtlichen Dolmetschfehler handelt. Außerdem werden noch leichte von schweren Formulierungsänderungen entschieden.

Zhao (2021: 95ff.) entwickelte die Fehleranalyse insbesondere aufbauend auf der Methode von Barik (1975) weiter. Zhao (2021) unterteilt die gedolmetschten Segmente anhand ihrer Syntax in Satzteile und untersucht diese auf ihre Genauigkeit bzw. Entsprechung mit dem Ausgangstext hin in einem schrittweisen Verfahren. Zunächst soll dabei für ein zielsprachliches Segment erkannt werden, ob dieses die Bedeutung des ausgangssprachlichen Segments ausdrückt. Falls ja, dann handelt es sich um eine Entsprechung. Falls nicht, dann soll das nicht mit dem Ausgangssegment übereinstimmende, daher ungenaue, Zielsegment dahingehend analysiert werden, ob es sich dabei um eine Hinzufügung handelt oder um eine Ersetzung. Im Falle einer unvollständigen Wiedergabe handelt es sich um eine Verkürzung bzw. Auslassung. In einem weiteren Schritt soll die Frage gestellt werden, inwiefern die Ungenauigkeit bzw. Unvollständigkeit die Bedeutung des ausgangssprachlichen Segments beeinflusst. Die Fehler werden dementsprechend in leichte, schwere und kritische Fehler eingeteilt (‚minor', ‚major' und ‚critical'). Die Analyseschritte wurden von Zhao (2021) übersichtlich graphisch dargestellt und werden auch hier zum besseren Verständnis in Abbildung 2 veranschaulicht.

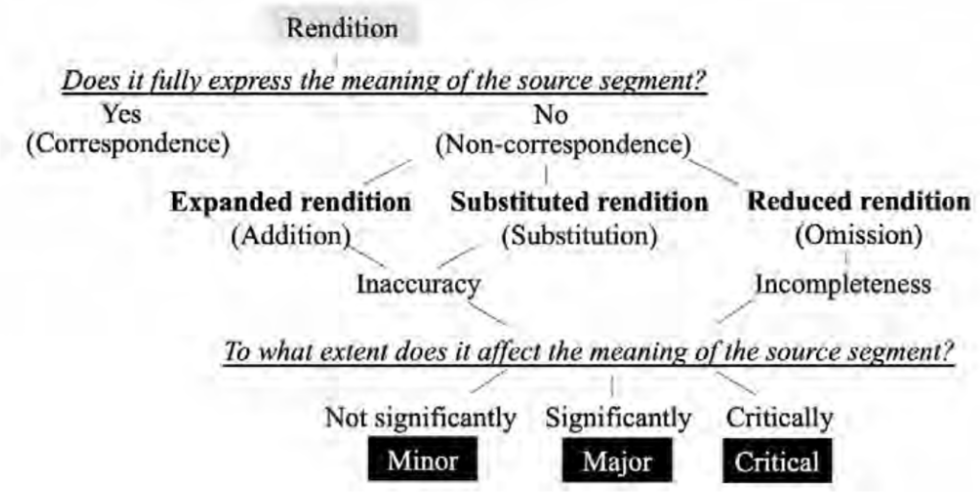


Abbildung 2: Bewertungsmodell nach Zhao (2021: 95)

Als Kritikpunkt zu der präsentierten Fehleranalyse nach Barik (1975) sowie folglich ihrer Weiterentwicklung nach Zhao (2021) kann gesagt werden, dass die Qualitätsbewertung lange Zeit auf einen Vergleich des Ausgangs- und Zieltexts beschränkt blieb, ohne den Einfluss eines gemeinsamen situativen Kontexts zu berücksichtigen (Pradas Macías, 2023: 69), welcher beispielsweise im Rahmen des Simultandolmetschens dank weiterentwickelter Modelle berücksichtigt wird (vgl. insbesondere Pöchhacker, 1994b).

Neben der Methode von Zhao (2021), die sich aufgrund einer stufenweisen Vorgehensweise für die Bewertung der Sinneinheiten eignet, lassen sich Problemauslöser separat betrachten und lediglich als durch die TN „verdolmetscht" oder „nicht verdolmetscht" bewerten. Die binäre Einteilung in die Kategorien verdolmetscht oder nicht verdolmetscht der analysierten Einheiten kann auch in der vorliegenden Arbeit gewählt werden, da Problemauslöser oft keine redundante Information enthalten und somit nicht umschrieben werden können, sondern lediglich „transkodiert" werden müssen (vgl. Mankauskiené 2018a: 16). Außerdem erlaubt diese Vorgehensweise eine punktuelle Bewertung an ausgewählten Stellen. Inspiriert von der Beurteilungsmethode nach Gieshoff & Albl-Mikasa (2022: 216) wird somit eine Bewertung in „verdolmetscht", repräsentiert durch (1), und „nicht verdolmetscht", repräsentiert durch (0), für das Verdolmetschen der Problemauslöser vorgenommen.

Die Kombination der Segmentierung nach Sinneinheiten und Problemauslösern erlaubt einerseits eine übergreifende, andererseits eine punktuelle Bewertung der Verdolmetschungen. Neben der beschriebenen Vorgehensweise gibt es zahlreiche weitere Vorgehensweisen. Pöchhacker (2016: 143) kritisiert im Zusammenhang mit weiterentwickelten und auf die Fehleranalyse nach Barik (1975) aufbauenden Modellen die Tatsache, dass Autor*innen im Zuge des Vergleichs der Ausgangs- und Zieltexte eigene Schemata verwenden, was wiederum eine Fülle an Klassifikationssystemen zur Konsequenz hat.

Die angesprochene Vielzahl an Analyseverfahren zur Kategorisierung und Bewertung der Genauigkeit zeigt sich nicht zuletzt dann, wenn Methoden für spezifische Elemente, wie Zahlen als Problemauslöser, entwickelt werden. Die Kategorisierung von Defrancq & Fantinuoli (2020: 86) bezieht sich aufgrund des im Rahmen der Evaluierung durchgeführten Experiments in erster Linie auf die Wiedergabe von Zahlen. Sie unterteilen in diesem Sinne gedolmetschte Segmente in vollständige Wiedergaben, Annäherungen, zusammenhängende Ersetzungen und unzusammenhängende Ersetzungen.

Auch Desmet et al. (2018: 21) verwenden für die Bewertung von gedolmetschten Zahlen eine Liste an Kategorien der Auslassungen, Annäherungen, lexikalischen Fehler, Vertauschungen, syntaktischen Fehler, phonologischen Fehler und anderen Fehler, wiederum in Bezug auf diese spezifischen Problemauslöser.

Obwohl die präsentierten Methoden von einer Vielfalt an unterschiedlichen Messungsverfahren der Genauigkeit zeugen, kann die jeweilige Methode beispielsweise dank der Bewertung durch (mehrere) professionelle Dolmetscher*innen sowie dank der Berechnung einer „inter-rater reliability", also ähnlicher Ergebnisse durch die jeweiligen Beurteilenden, um eine weitere Komponente ergänzt werden (vgl. Pöchhacker, 2016: 141).

Eine wesentliche Kritik von Gile (1992: 188) in Bezug auf Methoden zur Bewertung der Genauigkeit umfasst schlussendlich, dass linguistische Äquivalenz aufgrund unterschiedlicher Bedeutungen unmöglich oder kontraproduktiv sei und dass ein akzeptabler Zieltext eine gewisse Form von Abweichung erfordert. Somit müssten für eine umfassende Beurteilung der Darbietung zahlreiche weitere Parameter berücksichtigt werden, nicht zuletzt prosodische Merkmale wie etwa die Flüssigkeit und die Intonation (vgl. Pöchhacker, 2016: 214f.), ebenso wie mögliche idiomatische, folglich von der sogenannten Äquivalenz abweichende, oder auch paraverbale Elemente (vgl. ‚pragmatic impact', Pöchhacker, 2016: 148), die es ebenfalls für eine dahingehende Analyse einzugrenzen und zu definieren gilt.

Dabei darf nicht vergessen werden, dass die Bewertung einer Verdolmetschung nicht allein auf Parametern der Genauigkeit basieren kann. Es sei hier kritisch angemerkt, dass das Erkennen einer qualitativen Verdolmetschung nicht allein durch das Zählen von Fehlern erkannt werden kann, da die Kriterien bzw. Fehlerkategorien, wie aus dem Abschnitt 2.1 ersichtlich ist, nicht immer die gleichen sind, sondern je nach der beurteilenden Person voneinander abweichen können.

Die gesamte Darbietung besteht aus zahlreichen weiteren Faktoren. Diese umfassen beispielsweise die Flüssigkeit, u. a. gemessen anhand der Zögerungen, die Stimmqualität, das (un)genaue Artikulieren von Wörtern, die Verwendung paraverbaler, also außersprachlicher, Elemente sowie die Schnelligkeit, wobei all diese Faktoren den allgemeinen Eindruck eines gesprochenen Textes formen können (Pöchhacker, 1994a: 236).

Moser-Mercer (1996: 46) rückt in der Beurteilung der Qualität neben der dolmetschenden Person mit ihren eigenen Vorstellungen auch die Nutzer*innen der Verdolmetschung mit ihren Bedürfnissen und Erwartungen in den Vordergrund, ebenso wie Auftraggeber*innen, Kolleg*innen, etc. Zudem kann eine Veränderung der Qualität bzw. ausgewählter

Aspekte, wie der Genauigkeit, anhand der Manipulation unabhängiger Variablen von Interesse sein (vgl. dazu auch Pöchhacker, 2016: 142).

Abschließend ist aufgrund der entstehenden Vielfalt an Analyseperspektiven und der damit einhergehenden Fülle an zu berücksichtigenden Faktoren ersichtlich, dass nicht *die* Qualität als übergreifendes Kriterium bestimmt werden kann. Pöchhacker (1994a: 242) spricht daher auch von „quality under the circumstances".

Nachdem nun auf verschiedene Analyseperspektiven der Qualitätsbestimmung von Verdolmetschungen eingegangen wurde, leitet das folgende Kapitel das übergreifende Thema des computergestützten Dolmetschens ein. Neben einer Erläuterung der Funktionsweise von künstlicher Intelligenz anhand der automatischen Spracherkennung und der neuronalen maschinellen Übersetzung soll dabei insbesondere auf den Forschungsstand zum computergestützten Simultan- und Konsekutivdolmetschen und die entwickelten Anwendungen eingegangen werden.

3. Computergestütztes Dolmetschen

Der Bericht des ALPAC (kurz für ‚Automatic Language Processing Advisory Committee') der National Academy of Sciences kann als Ursprung der Forderung nach einer Technologieunterstützung für Translator*innen angesehen werden. In den Empfehlungen wird festgehalten, dass die Computerlinguistik einen Beitrag zur Erforschung der maschinellen Translation leisten soll, nicht zuletzt auf empirische Art, um u. a. den Arbeitsprozess von Translator*innen zu beschleunigen und um die einzelnen Arbeitsphasen zu bewerten (ALPAC, 1966: 34).

Die Fähigkeit eines Computers, Sprache zu verstehen, wird in der erwähnten Disziplin der Computerlinguistik erforscht. Carstensen et al. (2010: 1) definieren die Computerlinguistik als das Fachgebiet, das sich mit der maschinellen Verarbeitung natürlicher Sprache beschäftigt. Dabei unterscheiden Carstensen et al. (2010: 2) diese wissenschaftliche Disziplin anhand von vier Aspekten. Erstens, als Teildisziplin der Linguistik, zweitens, als Grundlage für die Schaffung von Programmen zur Verarbeitung sprachlicher Daten, etwa mittels Sprachkorpora, drittens, als Umsetzung sprachlicher Phänomene auf dem Computer, was wiederum auch als maschinelle Sprachverarbeitung definiert wird, sowie viertens, im Zuge einer praxisorientierten Konzeption von Sprachsoftware. Das in der vorliegenden Arbeit herangezogene Tool Sight-Terp (vgl. Abschnitt 3.3.2) wäre daher in der Computerlinguistik im Bereich der maschinellen Sprachverarbeitung bzw. bei der anwendungsorientierten Software angesiedelt. Im Zusammenhang mit der Computerlinguistik bildete sich zunächst die computergestützte Übersetzung aus (‚computer-assisted translation', kurz CAT), die Übersetzer*innen bei ihren kognitiven Anstrengungen entlasten und somit dank einer Zusammenarbeit zwischen Mensch und Maschine neue Möglichkeiten eröffnen würde (Carl, 2021: 511).

Die Ursprünge der technologischen Integration beim Dolmetschen gehen auf die Entstehung des Simultandolmetschens zurück und üben seither einen Einfluss darauf aus, wie Dolmetschaufträge in den verschiedenen Phasen ihrer Ausführung durch die dolmetschende Person abgewickelt werden. Computergestütztes Dolmetschen (‚computer-assisted interpreting', kurz CAI) definiert Fantinuoli dabei als Computersoftware, die der dolmetschenden Person in der Verbesserung der Qualität und Produktivität helfen soll.

Pöchhacker (2016: 190) beschreibt zudem, dass das Simultandolmetschen bereits in seinen Ursprüngen stark von modernen Kommunikationstechnologien abhing und dass es sich in seiner prinzipiellen Umsetzung, nämlich per elektroakustischer Übertragung, seit seiner Entwicklung in den späten 1920er-Jahren kaum gewandelt hat. Eine Definition der AIIC, in Anlehnung an die ISO-Normen (‚International Organization for Standardization'), macht die

technologische Komponente in diesem Arbeitsmodus deutlich sichtbar. Demnach sollten Simultandolmetscher*innen in einer schalldichten Kabine ein Pult, das eine klare Unterscheidung zwischen dem ausgangssprachlichen Input und dem zielsprachlichen Output zulässt, eigene Kopfhörer mit lautstärkenverstellbarem Ton sowie ein Mikrofon zur Verfügung stehen (AIIC, 2023).

Neben dem Arbeiten in Dolmetschkabinen wird das computergestützte Dolmetschen als Mittel eingesetzt, um die Arbeitsschritte im Dolmetschprozess weiter zu rationalisieren und zu optimieren (Fantinuoli, 2017: 3). Als typisches Beispiel für eine Kosten- und Zeitoptimierung sei das simultane Ferndolmetschen (‚remote simultaneous interpreting', kurz RSI), etwa in Form webbasierter Videokonferenzen, angeführt. Dieses ermöglicht es der dolmetschenden Person, nicht mehr bei dem zu verdolmetschenden Ereignis physisch anwesend zu sein (Pöchhacker, 2016: 193). RSI ist somit ein Beispiel dafür, dass die technologischen Fortschritte beim Dolmetschen zur Schaffung neuer Dolmetschmodi und zu einer erweiterten Reichweite der angebotenen Dienste beitragen (Chen & Kruger, 2022: 399).

Braun (2006: 5) unterscheidet beim Ferndolmetschen Verdolmetschungen im Zuge kommunikativer Anlässe, bei denen die Hauptteilnehmenden selbst an verschiedenen Orten sind, wie etwa bei einer Telefonkonferenz in Form von Audio- bzw. Videokonferenzen, von Verdolmetschungen im Rahmen von Veranstaltungen, bei denen sich die Hauptteilnehmenden am selben Ort treffen und bei dem nur die dolmetschende Person zugeschaltet wird. Das Video-Ferndolmetschen im konsekutiven Modus (‚video remote interpreting', kurz VRI) wird hingegen v. a. in Bereichen wie dem Gesundheitswesen eingesetzt (Pöchhacker, 2016: 22).

Fantinuoli (2018a: 155f.) unterteilt nun die beim Dolmetschen herangezogenen Informations- und Kommunikationstechnologien (‚information and communication technologies', kurz ICT) je nachdem, wie sie mit der dolmetschenden Person bzw. mit dem Dolmetschauftrag interagieren. Somit werden die prozessorientierten von den situationsorientierten Technologien unterschieden (‚process-oriented' und ‚setting-oriented'). Erstere umfassen Technologien mit unterstützenden Funktionen, die die dolmetschende Person in den verschiedenen Phasen eines Dolmetschauftrags, also vor, während bzw. danach, beispielsweise dank Terminologieverwaltungssystemen oder Korpusanalyseanwendungen, begleiten sollen. Diese seien unabhängig vom Dolmetschmodus zu betrachten, wobei Fantinuoli (2018a: 155f.) anmerkt, dass sie einen direkten Einfluss auf die kognitiven Prozesse beim Dolmetschen haben könnten. Prozessorientierte Technologien seien zudem ein Bestandteil des computergestützten Dolmetschens. Die in der vorliegenden Arbeit herangezogene Software Sight-Terp kann

folglich als eine prozessorientierte Technologie eingestuft werden, da sie den Dolmetscher*innen bei der Verbesserung ihrer Dolmetschleistung dienen soll (vgl. Ünlü, 2023: 5f.).

Unter situationsorientierten Technologien versteht Fantinuoli (2018a: 155f.) dann jene Anwendungen, die den Dolmetschprozess an sich umgeben, wobei diese Kategorie die benötigten Geräte zum Ferndolmetschen oder auch Lernplattformen umfasst. Letztere haben insbesondere einen Einfluss auf die äußeren Bedingungen, unter denen die Verdolmetschung erfolgt. Daher hätten sie kaum einen Einfluss auf die kognitiven Prozesse beim Dolmetschen. Situationsorientierte Technologien waren u. a. für die Entwicklung des Simultandolmetschens zentral und würden auch künftig das Dolmetschen beeinflussen.

Eine weitere Kategorisierung kann dahingehend getroffen werden, wofür die Anwendungen zum computergestützten Dolmetschen konkret eingesetzt werden. Corpas Pastor (2021a: 41) unterteilt diese in drei Kategorien, nämlich zur Verwendung für das Terminologiemanagement, für das Notieren im digitalen konsekutiven Arbeitsmodus (vgl. digital notepad, Abschnitt 1.2.2) sowie zur Nutzung von Speech-to-Text-Technologie (kurz S2T), also betreffend die Umwandlung lautsprachlicher in schriftliche Segmente (vgl. Abschnitt 3.1.1).

Ortiz & Cavallo (2018: 13f.) bieten eine Liste an CAI-Anwendungen und ordnen diese anders gewählten Kategorien zu, nämlich der Ausbildung, der Vorbereitungsphase und der Terminologieverwaltung. Außerdem unterscheiden sie in ihrem weiteren Werk neben einer grundlegenden Einteilung in Anwendungen für Übersetzer*innen im Vergleich zu Anwendungen für Dolmetscher*innen dann zusätzlich aufgrund der beiden Arbeitsmodi simultan oder konsekutiv einige Anwendungen, die jeweils nur in einem der beiden Modi bzw. für spezielle Zwecke, wie für Korpora oder zur Glossarerstellung, verwendet werden können. Dadurch ergibt sich eine Auflistung von Tools, die das computergestützte Dolmetschen anhand der Bezeichnung, der Hauptfunktion, wie Glossarverwaltung, Korpuserstellung, Ausbildungsmaterial etc. überblicksmäßig darstellt (Ortiz & Cavallo, 2018: 17f.). Ein Beispiel für eine zu Ausbildungszwecken entwickelte Anwendung des computergestützten Dolmetschens entstammt einer Studie von Gran et al. (2002), in deren Rahmen zwei computerbasierte interaktive Projekte entwickelt wurden, die das Lehren des Konferenzdolmetschens erleichtern sollen.

Wiederum eine andere Möglichkeit zur Kategorisierung von CAI-Tools zeigen Guo et al. (2023: 92), da sie die Anwendungen nicht aufgrund ihres Verwendungszwecks einteilen, sondern aufgrund ihres Designs. In Abbildung 3 werden neben den für bestimmte Zwecke eingeteilten Technologien, wie jene für das Ferndolmetschen, für die maschinelle Verdolmetschung, etc. die CAI-Tools als eigene Kategorie präsentiert, die sich wiederum aufteilen lässt

in sogenannte „first-generation“ und „second-generation tools“, also die ersten vorhandenen Anwendungen und die darauffolgenden. Diese Einteilung erinnert an jene von Fantinuoli (2018a: 164f.) im speziellen Kontext des Simultandolmetschens, jedoch werden hier auch die künftig relevanten „next-generation tools“ als weitere Unterkategorie angeführt. Guo et al. (2023: 93) erkennen, dass Dolmetscher*innen zunehmend mehr von „second-generation tools“ Gebrauch machen sowie dass diese zwar bis zu einem gewissen Grad die kognitiven Kapazitäten erweitern können, jedoch häufig (noch) nicht Anwendung finden.

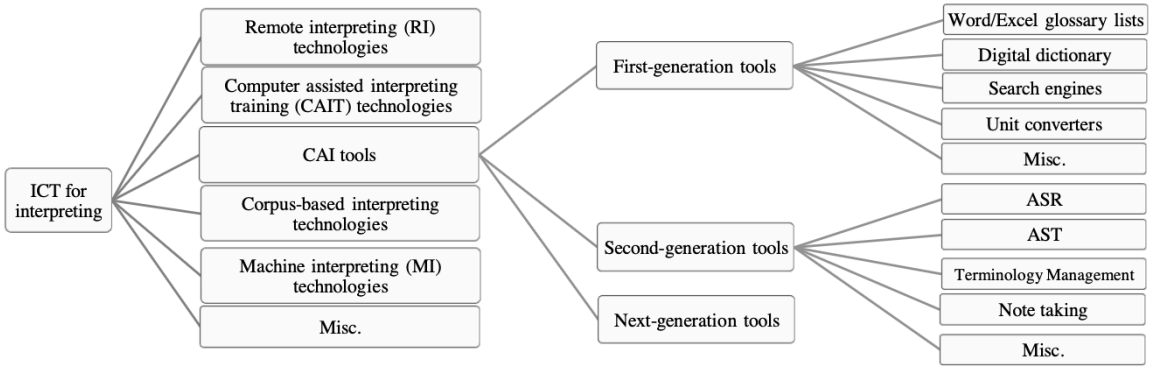


Abbildung 3: Kategorisierung der CAI-Tools (Guo et al., 2023: 93)

Aufgrund des vielversprechenden Potenzials spricht Fantinuoli (2018a: 170) von der Notwendigkeit, sowohl auf der prozess- als auch auf der produktorientierten Ebene Forschungsarbeit zum computergestützten Dolmetschen zu leisten, um die Vor- und Nachteile von CAI-Tools zu verstehen. Im Rahmen der prozessorientierten, experimentellen Forschung könnten Techniken wie etwa die Blickerfassung (‚eye-tracking‘) aufschlussreiche Ergebnisse in Hinblick auf den tatsächlichen Gebrauch der Tools beim Dolmetschen liefern können. Die produktorientierte Forschungsebene nach Fantinuoli (2018a: 167f.) umfasst die Analyse gedolmetschter Texte, die allgemein komparativer, also vergleichender, Natur ist. Letztere Forschungsebene wurde auch in der vorliegenden Arbeit gewählt (vgl. Kapitel 4).

Der folgende Abschnitt über künstliche Intelligenz soll einen Einblick in das computergestützte Dolmetschen liefern. Wesentliche Bestandteile stellen dabei die automatische Spracherkennung und die maschinelle Übersetzung dar, weshalb auf diese genauer eingegangen werden soll.

3.1 Künstliche Intelligenz beim Dolmetschen

Die Forschung rund um künstliche Intelligenz (‚artificial intelligence', kurz AI oder KI) begann mit dem Versuch, eine Maschine dazu zu bringen, Sprache zu nutzen, Abstraktionen und Konzepte zu schaffen und mit der Forderung, dass diese jegliche Probleme lösen könne, die bislang Menschen vorbehalten waren, sowie dass sich diese Maschine auch selbst verbessern kann (McCarthy et al., 2006: 12). Die Grundannahme ist dabei also, dass Maschinen lernen können, und zu diesem Zweck basiert die Technologie künstlicher Intelligenz auf „predictive models", die u. a. aufgrund statistischer Methoden und Algorithmen Entscheidungen treffen können (Ahmed, 2022: 334).

Der Einsatz von künstlicher Intelligenz im Bereich der Translationswissenschaft ist in seinem Ursprung stark mit der maschinellen Übersetzung (kurz MÜ) verknüpft (Carl, 2021: 500). Dabei basierte die MÜ ab den späten 1950er-Jahren auf der sogenannten „Physical Symbol Systems Hypothesis" (kurz PSSH), welche davon ausgeht, dass intelligente Computeranwendungen menschliche kognitive Prozesse repräsentieren können (Carl, 2021: 511).

Künstliche neuronale Netzwerke bilden in diesem Sinne die Basis für diese computergestützte Modellierung kognitiver Prozesse und versuchen die Arbeitsweise des Gehirns nachzubilden (Carl, 2021: 504). Carl (2021: 504) erwähnt als einen großen Kritikpunkt an dieser Methode, dass das Gedächtnis in der Realität viel mit Symbolen, Metaphern und Kombinationen der geistigen Darstellungen arbeitet, welche wiederum schwer per neuronaler Netzwerke sowie ohne Informationsverlust abzubilden seien.

Daran anschließend lässt sich hinterfragen, inwiefern künstliche Intelligenz für das Übersetzen und Dolmetschen eines zugrundliegenden Gedankens in Hinblick auf die „théorie du sens" verwendet werden kann (vgl. Abschnitte 1.2.2 und 2.2.2).

Zur Integration von künstlicher Intelligenz beim Dolmetschen werden daher aufgrund der technologischen Fortschritte und nicht zuletzt dank der Forschung zum computergestützten Konsekutivdolmetschen von Chen & Kruger (2022) sowie auch der Funktionalitäten von Sight-Terp (Ünlü, 2023), die im empirischen Teil dieser Arbeit zur Anwendung kommen, die automatische Spracherkennung sowie die maschinelle Übersetzung als deren wesentliche Bestandteile angeführt und in den folgenden Abschnitten vorgestellt.

3.1.1 Automatische Spracherkennung

Die automatische Spracherkennung (‚automatic speech recognition', kurz ASR) wandelt Sprachsignale, also ausgesprochene Wortsequenzen, mittels Algorithmen, die Teile einer Software bilden, in eine textuelle Darstellung um (Besacier et al., 2014: 89). Automatisch bedeutet hier, dass der Vorgang der Transkription (zunächst) ohne menschliches Eingreifen erfolgt (vgl. Chen & Kruger, 2022: 402).

ASR basiert nun meist auf einem verborgenen Markov-Modell (‚Hidden Markov Model'), welches die erkannten akustischen Signale mit Einträgen im Lexikon und mit einer syntaktischen Überprüfung abgleicht, die trotz großer Fortschritte noch vor großen Herausforderungen steht, was das Zusammenspiel akustischer und sprachlicher Modelle in einem einzigen System betrifft (Pöchhacker 2016: 197).

Die automatische maschinelle Verdolmetschung, die dank einer Speech-to-Text- (‚S2T'), oder auch Speech-to-Speech-Technologie ermöglicht wird, also per Umwandlung von sprachlichen Segmenten in schriftliche oder lautsprachliche Texte, ist laut Chen & Kruger (2022: 404) ebenfalls noch weit davon entfernt, zufriedenstellende Ergebnisse zu liefern und bedarf daher ausgiebiger Forschung im Bereich der computergestützten menschlichen Verdolmetschung sowie der menschengestützten maschinellen Verdolmetschung (‚computer-assisted human interpreting' und ‚human-assisted machine interpreting').

Die Anwendung von automatischer Spracherkennung erlaubt es den Dolmetscher*innen dank automatischer Transkription von potenziellen Problemstellen, wie Zahlen, Abkürzungen und Eigennamen, im konsekutiven Modus Segmente des Gesagten vom Blatt zu übersetzen, um eine genauere und vollständigere Verdolmetschung zu ermöglichen (Fantinuoli 2017: 2; vgl. dazu auch Abschnitt 2.2).

Cheung & Li (2022: 5) treffen eine klare Unterscheidung im simultanen Arbeitsmodus, nämlich zwischen Simultandolmetschen mit Text, welches ein Manuskript einer vorzutragenden Rede beinhaltet, und Simultandolmetschen mit automatischer Spracherkennung, bei der dank ASR eine textuelle Darstellung mündlichen Ausdrucks zum Tragen kommt. Das bedeutet, dass u. a. Wortwiederholungen, unvollständige Sätze und Füllwörter in dem mittels ASR erstellten Transkript vorkommen können. Das macht die mithilfe der automatischen Spracherkennung erstellte Rede zwar möglicherweise unübersichtlicher, aber wesentlich genauer, da sie dem tatsächlich geäußerten Original entspricht.

Cheung & Li (2022: 5) erwähnen, dass beim Simultandolmetschen mit Text dieser als Gesamtheit zur Verfügung steht und daher auch gesamtheitlich analysiert werden kann, was

wiederum bei einer Erstellung automatischer Untertitel mittels ASR nicht möglich ist, da nur Segmente nach und nach angezeigt werden. In Bezug auf das Konsekutivdolmetschen mit einem von ASR erstellten Text sei im Gegensatz dazu angemerkt, dass dieser bereits vor Beginn der Verdolmetschung als komplette Version des Gesagten zur Verfügung steht und somit ggf. auf Makroebene analysiert werden kann.

Um die oben erwähnten Herausforderungen der automatischen Spracherkennung zu überwinden, ist es u. a. notwendig, dass Letztere gewisse Anforderungen erfüllt, beispielsweise, dass die ASR de facto sprecherunabhängig ist, dass es eine kurze Reaktionszeit auf die Äußerung gibt und dass Fachterminologie genau erkannt wird (Fantinuoli 2017: 2). Doch die sprachlichen Signale sind nicht die einzigen Komponenten einer Äußerung, die sich in einer Verdolmetschung widerspiegeln sollten.

Schlussendlich erwähnt Fantinuoli (2017: 4) hierzu, dass neben den akustischen Signalen einer Rede auch zahlreiche weitere Informationen, wie u. a. jene über die Vortragenden selbst oder das eigene Welt- und Fachwissen beim reinen Verlassen auf die technischen Anwendungen verloren gehen. Diese Informationen sind aber in der Verdolmetschung sehr wohl entscheidend, nicht zuletzt um beispielsweise Fehlendes zu ergänzen.

Neuronale Netzwerke und tiefes Lernen (‚Deep Learning') haben die Qualität der automatischen Spracherkennung erheblich verbessert (Fantinuoli 2017: 4) und könnten einen möglichen Lösungsansatz für einen Teil der beschriebenen Problematik liefern. Auch die in dieser Arbeit untersuchte Anwendung Sight-Terp verwendet ein System zur maschinellen Übersetzung, ebenso wie eine automatischen Spracherkennung, welches auf einer Modellierung tiefgreifender neuronaler Netzwerke basiert (Ünlü, 2023: 67).

3.1.2 Neuronale maschinelle Übersetzung

Maschinelle Übersetzung wird als die automatische Übersetzung eines Ausgangstexts in einen Zieltext definiert (Costa-jussà, 2018: 947). In den letzten Jahrzehnten gab es dabei einen Wandel von der regelbasierten über die statistische bis hin zur neuronalen maschinellen Übersetzung, wobei Dolmetscher*innen von den diesbezüglich gemachten technologischen Fortschritten bislang weniger profitierten als Übersetzer*innen (Chen & Kruger, 2022: 403).

Die neuronale maschinelle Übersetzung (‚neural machine translation', kurz NMT) hat ihren Ursprung im Jahr 2014 und wurde durch eine erhöhte Verarbeitungskapazität der Computer sowie eine Vektorenmethode ermöglicht (Carl, 2021: 505). Letztere sammelt semantische Ähnlichkeiten in großen Korpora, bei denen Wörter und Strukturen in ähnlichen Kontexten und mit ähnlichen Bedeutungen vorkommen. Neuronale maschinelle Übersetzung wird

dadurch ermöglicht, dass große Datenmengen mittels verknüpfter künstlicher Neuronen, die den biologischen neuronalen Netzwerken nachempfunden sind, in das System eingespeist werden (Luo, 2018: 1).

Das sogenannte tiefe Lernen (‚deep learning') ist dabei eine Weiterentwicklung neuronaler Netzwerke, die aus mehrschichtigen Modellen bestehen und aufgrund zahlreicher Beispiele und ohne programmiert zu werden selbständig Aufgaben erlernen können, wobei die maschinelle Übersetzung nur eine von vielen Bereichen ist, bei denen dieses Konzept Anwendung findet (Costa-jussà, 2018: 947). Costa-jussà (2018: 958) beschreibt in diesem Zusammenhang, dass die neuronalen Netzwerke der neuronalen maschinellen Übersetzung dahingehend trainiert werden, einen Zieltext unter Vorhandensein eines Ausgangstexts vorherzusagen (vgl. ‚predicitive models', Abschnitt 3.1), und erläutert die zur Anwendung kommenden statistischen Wahrscheinlichkeitsberechnungen.

Luo (2018: 1) merkt hingegen kritisch an, dass maschinelle Übersetzungsprogramme jedes Wort und jedes Konzept jeder Sprache sammeln müssten, um bei dieser Aufgabe menschliche Übersetzer*innen in ihrer Arbeit zu übertreffen, sowie dass es für die neuronale maschinelle Übersetzung unmöglich sei, dieser Aufgabe gewachsen zu sein. Luo (2018: 1) nennt linguistische Kreativität und Unvorhersehbarkeit als einige Gründe dafür.

Ein weiterer Kritikpunkt in Bezug auf die neuronale maschinelle Übersetzung liegt darin, dass die vom Computer gemachten Fehler immer mehr jenen eines Menschen ähneln und daher schwer zu entdecken sind bzw. oft nicht mehr von einem Menschen ausgebessert werden (Carl, 2021: 513).

In diesem Zusammenhang ist das Aufkommen von postmaschineller Bearbeitung (‚post editing', kurz PE) zu erwähnen. Luo (2018: 1) erkennt dabei, dass die maschinelle Übersetzung es den Translator*innen erlaubt, mehr Zeit mit dem Wesentlichen zu verbringen, sowie auch dass sich die Arbeit in manchen Bereichen von den Übersetzer*innen hin zu den Post-Editor*innen bewegen wird. Das Wesentliche könnte also die Korrektur der maschinell erstellten Übersetzung bedeuten. Interessanterweise hat die Nachbearbeitung bereits die Erwartungen der Nutzer*innen verändert, da diese vor allem Translate, die „fit for use", also anwendungsgerecht sind, bevorzugen würden (vgl. Yamada, 2014: 50).

Diese Erkenntnis wirft neben geänderten Arbeitsbedingungen für Translator*innen auch weitere Fragen auf, die die Qualität (vgl. Kapitel 2) der mittels neuronaler Netzwerke produzierten Translate in Hinblick auf die geänderten Anforderungen der Nutzer*innen betrifft, da die maschinelle Übersetzung aufgrund ihres kurzen Daseins und aufgrund der

Tatsache, dass sie abhängig vom Sprachenpaar und der Menge an verfügbaren Daten Schwankungen aufweist, noch großen Verbesserungsbedarf hat (vgl. Costa-jussà, 2018: 960).

Die Anwendung der neuronalen maschinellen Übersetzung beim Dolmetschen ist nun beispielsweise bei der Speech-to-Speech Übersetzung vorzufinden, bei der die Maschine vollautomatisch, das heißt ohne Eingreifen durch eine dolmetschende Person, von einer Ausgangs- in eine Zielsprache „verdolmetscht". Diese Funktionsweise ist in ihrer Grundform für den Input mit der oben beschriebenen automatischen Spracherkennung verknüpft und setzt beim Output auf eine sprachliche Synthese (vgl. Pöchhacker, 2016: 198).

So wurde das System Verbmobil entwickelt, um in den Sprachen Deutsch, Englisch und Japanisch sowie in drei ausgewählten Einsatzbereichen die Nutzer*innen der Anwendung in ihrer mündlichen Konversation zu begleiten (Wahlster, 2000: 3ff.). Man denke hier an heutige Anwendungen auf mobilen Endgeräten wie etwa Google Assistant, bei dem der Modus Dolmetschen ausgewählt werden kann und in einer Konversation zwischen zwei Sprachen etwa auf einem Mobilgerät unkompliziert verwendet werden kann.

Sight-Terp kombiniert in diesem Sinne beide der beschriebenen Funktionalitäten künstlicher Intelligenz, d. h. die automatische Spracherkennung und die maschinelle Übersetzung, zur Anwendung für menschliche Dolmetscher*innen im konsekutiven Arbeitsmodus.

Die Fortschritte im Bereich der maschinellen Übersetzung führen somit einerseits zu einer Transformation des Berufs der Translator*innen (Luo, 2018: 1) bzw. zu einer Änderung des „Status des Berufs von Translator*innen" (O'Brien, 2021: 380). Andererseits entwickeln viele professionelle Dolmetscher*innen eine ablehnende Haltung gegenüber den prozessorientierten Technologien (vgl. Kapitel 3), da sie Zweifel am Wert ihrer intellektuellen Arbeit hervorrufen könnten (Fantinuoli, 2018a: 163).

Die folgenden Abschnitte unterteilen das computergestützte Dolmetschen erneut in die beiden Arbeitsmodi simultan und konsekutiv und zeigen relevante Anwendungen sowie den Forschungsstand auf.

3.2 Computergestütztes Simultandolmetschen

Computergestütztes Simultandolmetschen wurde zunächst in Form von Terminologie- und Glossarverwaltungssystemen entwickelt bzw. erforscht. Fantinuoli (2018a: 164) nennt diese auch „first-generation CAI tools", also die allerersten verfügbaren Anwendungen für das computergestützte Dolmetschen. Diese Tools ähneln Programmen wie Microsoft Word oder Excel und standen am Beginn der Integration von Technologie in diesem Bereich.

Fantinuoli (2009: 412) spricht im weiteren Verlauf von einer immer größer werdenden Forderung nach der Weiterentwicklung computerbasierter Anwendungen, die die Vorbereitung und die Nachbereitung eines Dolmetschauftrags unterstützen sollen. Um dieser Forderung gerecht zu werden, kam es zur Entwicklung der sogenannten „second-generation tools" (Fantinuoli, 2018a: 165).

Einem Überblick über die diesbezüglich verfügbaren Anwendungen dient das VIP-System (kurz für ‚**V**oice-text **I**ntegrated system for inter**P**reters'), welches online frei zur Verfügung steht und sowohl in der Dolmetschpraxis als auch in der Lehre als Unterstützung fungieren soll (Corpas Pastor, 2021b: 97). Es beinhaltet einen offenen Katalog von Werkzeugen und Ressourcen sowie plattformintegrierte Funktionen, die die Dolmetscher*innen in allen Phasen des Auftrags in Anspruch nehmen können (Corpas Pastor, 2021b: 99f.). Der offene Katalog enthält eine Liste an Tools für das Terminologiemanagement, wie InterpretBank, jene zum Notieren, wie Livescribe, sowie zahlreiche weitere Tools für die Speech-to-Text-Funktion und für das Ferndolmetschen. Dadurch wird ein Überblick und somit ein Vergleich der jeweiligen Anwendungen ermöglicht. Zur Hilfe bei der Vorbereitung sei auch die (halb-)automatische Korpuserstellungsfunktion von VIP angeführt (Corpas Pastor, 2021b: 109). Außerdem erkennt VIP sogenannte Named Entities (vgl. dazu Abschnitt 2.2.1), also Organisationen, Eigennamen, Orte, usw. und kann diese Information aus einem Text filtern (Corpas Pastor, 2021b: 115).

Im Folgenden werden einige der darin vorkommenden grundlegenden Anwendungen für das computergestützte Simultandolmetschen skizziert.

3.2.1 Interpreter's Help

Eines der bekanntesten Beispiele der CAI-Anwendungen erster Generation ist Interpreter's Help bzw. Boothmate[1], bei der mehrsprachige Glossare, ähnlich einer Excel-Liste, verwaltet werden können sowie in die bzw. aus der Anwendung im- bzw. exportierbar sind.

Die erkannte Terminologie kann manuell nebeneinander aufgelistet bzw. auch direkt aus einem vorhandenen Text extrahiert werden. Das Glossar hat daneben weitere Funktionalitäten. Beispielsweise kann ein erstelltes Glossar mit anderen Kolleg*innen geteilt und von diesen mitbearbeitet werden. Neben der Vorbereitung auf einen Dolmetscheinsatz dient Boothmate aufgrund der einfachen Anwendungsweise der Verwendung für ad hoc Suchabfragen beim Simultandolmetschen in der Kabine (vgl. Fantinuoli, 2018a: 164f.).

[1] abrufbar unter https://boothmate.app

3.2.2 InterpretBank

InterpretBank ist eine im Rahmen einer Dissertation entwickelte Anwendung mit vier Funktionalitäten, die Konferenzdolmetscher*innen vor, während und nach der zu dolmetschenden Veranstaltung unterstützen soll.

Die Funktionen werden in Module eingeteilt und umfassen ConferenceMode zur Verwendung während des Simultandolmetschens in der Kabine, TermMode für die Terminologiearbeit, CorpusMode für die Termextraktion und DrillMode für das Lernen von Glossaren (Fantinuoli, 2009: 413). Dank dieser Funktionalitäten sind die Anwendungsmöglichkeiten vielfältig, wie etwa das Sammeln von Korpora und die Extraktion linguistischer Informationen als Vorbereitung auf den Auftrag, das Erstellen und Verwalten von auftragsspezifischen Glossaren sowie das schnelle Abrufen von Terminologie vor und auch während der Verdolmetschung (Fantinuoli, 2016: 45ff.).

InterpretBank wurde zu einem späteren Zeitpunkt unter Verwendung von automatischer Spracherkennung weiterentwickelt. Dabei schuf Fantinuoli (2017: 7) einen Prototyp, der aufbauend auf InterpretBank und unter Verwendung sprachunabhängiger Algorithmen den Dolmetscher*innen automatische Transkriptionen der lautsprachlich geäußerten Botschaft liefert sowie terminologische Einträge und Zahlen zur Verfügung stellt.

3.2.3 SmarTerp

Ein weiteres System für das computergestützte Simultandolmetschen ist SmarTerp (Rodríguez et al., 2021: 102ff.), welches eine automatische Audiotranskription sowie die Markierung wichtiger Elemente, wie Named Entities und kognitiv anspruchsvolle Segmente (vgl. Problemauslöser, Abschnitt 2.2.1), in einer Anwendung kombiniert. SmarTerp soll vor allem beim simultanen Ferndolmetschen (‚remote simultaneous interpreting', kurz RSI) der Unterstützung dienen und über den gesamten Arbeitsablauf der Ausführung eines Dolmetschauftrags hinweg verwendet werden.

3.2.4 KUDO Interpreting Assistant

Der KUDO Interpreting Assistant (Fantinuoli et al., 2022) zählt auch zu den grundlegenden CAI-Tools, welche für den simultanen Arbeitsmodus entwickelt wurden. Dieses System umfasst die beiden Hauptfunktionen der automatischen Glossarerstellung sowie eine Vorschlagsfunktion in Echtzeit für Terminologie, Zahlen und Eigennamen.

3.2.5 Forschungsstand

Neben der Entwicklung von Tools wurde das computergestützte Simultandolmetschen empirisch erforscht, oftmals im Kontext von Zahlen als Problemauslöser (vgl. Abschnitt 2.2.1).

Desmet et al. (2018: 18) führten im Bereich des computergestützten Simultandolmetschens ein Experiment durch, bei dem überprüft wurde, ob die Verfügbarkeit von technologischer Unterstützung die Genauigkeit der Wiedergabe von Zahlen in der Verdolmetschung verbessern kann. Dafür verwendeten sie ein System, das in der Ausgangsrede automatisch Zahlen erkennen kann und diese dann auf dem Bildschirm für die Dolmetscher*innen erscheinen lässt (Desmet et al., 2018: 25). In ihrer Studie fanden sie heraus, dass das Vorhandensein technologischer Unterstützung die Häufigkeit der korrekten Zahlenwiedergabe durch die dolmetschende Person erheblich steigerte, da Fehler um zwei Drittel reduziert werden konnten (Desmet et al., 2018: 25).

Ein weiteres Experiment wurde mit InterpretBank ASR von Defrancq & Fantinuoli (2020) durchgeführt, bei dem die Teilnehmenden mit und ohne Unterstützung des Tools simultan Reden dolmetschten und die Anwendung auch beurteilten. Dabei fanden Defrancq & Fantinuoli (2020: 97) heraus, dass die Genauigkeit der Zahlenwiedergabe in der computergestützten Verdolmetschung erhöht wurde. Jedoch erkannten sie auch, dass die Neuheit von InterpretBank ASR und folglich die Tatsache, dass die Dolmetscher*innen die Arbeit mit dem Tool nicht gewöhnt waren, zu einer erhöhten kognitiven Belastung führten. Dabei wurde zwar erkannt, dass das Vorhandensein der Anwendung Stress reduzieren kann, da eine Art Absicherung vorhanden ist, andererseits wurde kritisch angemerkt, dass jene Teilnehmenden, die sich zu sehr auf die Anwendung verließen, bei einem plötzlichen Wegfallen der Unterstützung auch einen Leistungseinbruch erlitten (Defrancq & Fantinuoli, 2020: 98).

In einer weiteren Studie untersuchten Pisani & Fantinuoli (2021: 187), welchen Einfluss der Einsatz von automatischer Spracherkennung auf die Verdolmetschung von Zahlen ausübte und inwiefern die Teilnehmenden von dieser Technologie profitieren würden. Pisani & Fantinuoli (2021: 194f.) konnten zeigen, dass ASR die Leistung erheblich verbessert.

Fantinuoli & Prandi (2021: 3) führten wiederum eine Studie durch, bei der die Leistung maschineller Verdolmetschungen mit der Speech-to-Text-Technologie anhand einer Messung der Genauigkeit an einem „gold standard“ (vgl. dazu Setton & Motta, 2007: 203 bzw. Abschnitt 2.2) mit den Verdolmetschungen durch menschliche Simultandolmetscher*innen verglichen wurden und beabsichtigten damit in erster Linie einen Rahmen für die Bewertung von Speech-to-Text- und Speech-to-Speech-Technologien zu schaffen (2021: 8).

Cheung & Li (2022: 7) führten ein Experiment durch, bei dem Videoaufnahmen in vier Abschnitte aufgeteilt wurden, wobei nur zwei von diesen Abschnitten automatisch generierte Untertitel mittels Spracherkennung enthielten. Die Teilnehmenden wurden gebeten, die Reden simultan zu dolmetschen. Dabei fanden Cheung & Li (2022: 14) heraus, dass die Genauigkeit der Verdolmetschungen verbessert wurde, aber dass sprachliche Interferenzen durch das Ablesen der Untertitel die Flüssigkeit der Darbietung beeinträchtigten. Außerdem fanden sie heraus, dass die Untertitel den Teilnehmenden ein Sicherheitsgefühl gaben.

Rodríguez González et al. (2023: 6) fanden in einer Studie heraus, dass die automatische Spracherkennung beim Ferndolmetschen ebenso Anwendung finden kann, da sie auch in diesem Arbeitsmodus zu genaueren und vollständigeren Verdolmetschungen führt.

Das bedeutet zusammenfassend, dass das Vorhandensein eines technologischen Hilfsmittels stets im Spannungsfeld zwischen einer nützlichen Absicherung und einer möglichen Ablenkung zu verorten ist, je nachdem, wie Dolmetscher*innen davon Gebrauch zu machen wissen, was wiederum mit der theoretischen und praktischen Erfahrung im computergestützten Arbeitsmodus zusammenhängt.

Die Kritik von Corpas Pastor (2021b: 96) umfasst in diesem Kontext, dass trotz der Nützlichkeit von CAI-Anwendungen, wie der automatischen Spracherkennung, die Vorteile aufgrund der bestehenden technischen Einschränkungen bei ASR und bei der sprachlichen Synthese begrenzt ausfallen. Ein Beispiel dieser technischen Defizite ist die Herausforderung der Entwicklung eines Tools zur automatischen Spracherkennung, welches Umstände, wie verschiedene Redestile, Hintergrundgeräusche und die Raumakustik berücksichtigen kann (vgl. Chen & Kruger, 2022: 402). Dieser Umstand zeigt die Grenzen der CAI-Tools.

3.3 Computergestütztes Konsekutivdolmetschen

Chen & Kruger (2022: 400) erkennen im Aufkommen des Technologieeinsatzes beim Konsekutivdolmetschen das Potenzial, die grundlegende Art seiner Durchführung neu zu formen und auch die kognitiven Prozesse zu beeinflussen. Sie forschten im Bereich des neu entstandenen Arbeitsmodus, des sogenannten computergestützten Konsekutivdolmetschens (‚computer-assisted consecutive interpreting', kurz CACI).

Das computergestützte Dolmetschen brachte bereits neue Arbeitsmodi hervor. Nicht zuletzt im Bereich des computergestützten Konsekutivdolmetschens gibt es auch weiteres Entwicklungspotenzial. Ein Forschungsziel von Chen & Kruger (2022: 416) bleibt es daher weiterhin, computergestützte Dolmetschmodi im konsekutiven bzw. simultanen Modus sowie beim Vom-Blatt-Übersetzen zu entwickeln.

Die Definition dieser Arbeitsmodi stellt sich jedoch aufgrund des Vorhandenseins mehrfacher Inputs als schwierig heraus. Beispielsweise werden bei Sight-Terp (vgl. Abschnitt 3.3.2) mehrere Arbeitsmodi bzw. -weisen zur Auswahl gestellt, z. B. das Vom-Blatt-Übersetzen des durch ASR automatisch erstellten ausgangssprachlichen Transkripts bzw. das mündliche Posteditieren der maschinellen Übersetzung in die Zielsprache bzw. auch die Nutzung computergestützt angefertigter bzw. klassischer Notizen (vgl. Ünlü, 2023: 66ff.). Die dolmetschende Person könnte dabei potenziell an ihre Belastungsgrenzen geraten.

Im Falle des SightConsec-Modus (vgl. Corpas Pastor, 2022: 96) sei an dieser Stelle angemerkt, dass trotz technischer Unterstützung, z. B. mittels automatisch angefertigter Transkription, ähnlich wie beim Simultandolmetschen mit Text, eine Abweichung vom geäußerten Original entstehen kann, nicht etwa weil die vortragende Person in ihrem Vortrag abweicht, sondern weil die Transkription eine potenzielle Fehlerquelle darstellen kann. Trotzdem wäre das tatsächlich Gesagte und nicht das Transkribierte zu verdolmetschen (vgl. Pöchhacker, 2016: 20). Das Überprüfen auf Korrektheit des transkribierten Textes könnte wiederum den kognitiven Aufwand bei dieser Art der konsekutiven Verdolmetschung erhöhen und die Arbeitsweise beeinflussen (vgl. kognitive Aspekte, Kapitel 1).

Im Folgenden werden Anwendungen des computergestützten Konsekutivdolmetschens sowie der diesbezügliche Forschungsstand umrissen.

3.3.1 Smartpen

Zur technologischen Integration beim Konsekutivdolmetschen sei zunächst die Studie von Orlando (2014: 42f.) angeführt, in deren Rahmen die Anwendung der Technologie des digitalen Stifts Livescribe Smartpen Modell Pulse™ erprobt wurde. Der genannte Stift nimmt die Ausgangsrede auf und lässt eine nachträgliche Wiedergabe beim Antippen des jeweiligen notierten Segments zu. Somit ergibt sich als Modus eine Mischform, das sogenannte „Consec simul with notes“, also das konsekutive Simultandolmetschen mit Notizen.

Die Teilnehmenden wurden im Rahmen der Studie darum gebeten, zwei konsekutive Verdolmetschungen durchzuführen, nämlich einmal mit der traditionellen Methode, d. h. mit Papier und Stift, und einmal mit dem digitalen Stift. Die Ergebnisse der Studie weisen auf eine deutlich verbesserte Genauigkeit der Verdolmetschungen unter Verwendung des digitalen Stifts hin (Orlando, 2014: 51).

3.3.2 Sight-Terp

In diesem Abschnitt wird das im empirischen Teil zur Anwendung kommende Tool Sight-Terp genau erklärt. Sight-Terp ist eine webbasierte Software, die von Ünlü (2023: 65ff.) im Zuge einer Masterarbeit zum computergestützten Konsekutivdolmetschen, und insbesondere für die Anwendung im neu entwickelten Arbeitsmodus Sight-Consecutive, also konsekutives Dolmetschen in Kombination mit dem Vom-Blatt-Übersetzen (vgl. Corpas Pastor, 2022: 96), entwickelt wurde.

Sight-Terp ist eine webbasierte Anwendung. Um sie verwenden zu können, genügt es, lediglich im Browser die Webseite www.sightterp.net abzurufen und sich mit einer E-Mail-Adresse zu registrieren, denn die Nutzung ist kostenlos und frei zugänglich. Sobald man sich registriert hat, erscheint eine Benutzeroberfläche, die im oberen Teil zunächst das Logo bzw. den Namen von Sight-Terp enthält und dann darunter auf zwei Spalten aufgeteilt ist, welche die im Folgenden beschriebenen Funktionalitäten beinhalten.

Die Ausgangs- und Zielsprachen können vor Beginn der Aufnahme per Drop-Down-Menü jeweils oberhalb der linken und rechten Spalte ausgewählt werden. Die Liste enthält zahlreiche Sprachen sowie auch Sprachvarietäten (wie u. a. amerikanisches, britisches und kanadisches Englisch) und kann somit womöglich genau(er) auf die Bedürfnisse der jeweiligen Vortragenden abgestimmt werden.

Sobald die automatische Spracherkennung (kurz ASR) mit einem Klick auf „Start Recognition“ gestartet wird, erscheint die Ausgangsrede mittels automatischer Spracherkennung in Echtzeit als Fließtext bzw. wird nach dem tatsächlich Gesagten Wort für Wort transkribiert. Dieser Vorgang wird auch als Speech-to-Text-Übersetzung bezeichnet und wird bei Sight-Terp dank des Programms Microsoft Speech Translation API ermöglicht (Ünlü, 2023: 67).

Auf der rechten Seite erfolgt dann etwas zeitverzögert und basierend auf der dank ASR erstellten Transkription eine automatische maschinelle Übersetzung in die vorher ausgewählte Zielsprache, auf die sich Dolmetscher*innen im konsekutiven Arbeitsmodus neben der Transkription in der Ausgangssprache ebenfalls beziehen können.

Die Bildschirmaufnahme in Abbildung 4 zeigt die Benutzeroberfläche zur Veranschaulichung und lässt erkennen, dass das grundlegende Erscheinungsbild bereits auf den ersten Blick klar und unkompliziert aufgebaut ist.

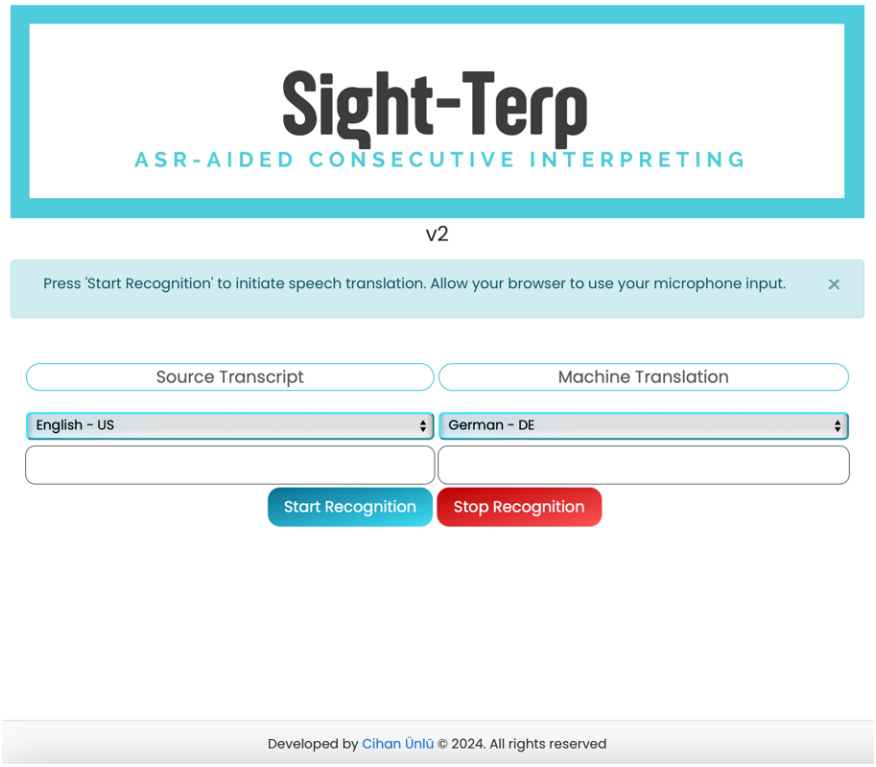


Abbildung 4: Benutzeroberfläche von Sight-Terp (www.sightterp.net, Datum des Zugriffs: 10.06.2024)

Die Transkription, die in der linken Spalte mit Beginn der Aufnahme erscheint, wird zudem zum besseren Überblick und im Sinne einer Nachempfindung der geforderten Struktur im Rahmen der Notizentechnik (vgl. Abschnitt 1.2.2) in nummerierte Absätze gegliedert. Außerdem werden Problemauslöser bzw. ein Teilaspekt, die sogenannten „Named Entities“ (vgl. Abschnitt 2.2.1), wie Namen, Organisationen, Orte, etc. ebenso wie Zahlen gelb markiert.

Die Segmentierung in Absätze erfolgt nach jeder längeren Pause von etwa zwei Sekunden (Ünlü, 2023: 69). Das System versucht dabei ebenfalls, dank der Ermittlung von Sprechpausen in einem Satz vorherzusehen, wann dieser zu Ende ist (Ünlü, 2023: 67).

Das bedeutet auch, dass schnellere Reden, die keine langen Pausen aufweisen, eventuell weniger gut segmentiert werden können, da das Programm das Ende eines alten und den Beginn eines neuen Gedankens schwer erkennen kann. Da die im durchgeführten Experiment der vorliegenden Arbeit verwendeten Redebeiträge (vgl. Abschnitt 4.2.1) allesamt ein rasches Tempo zwischen 150 bis 160 Wörtern pro Minute aufwiesen (vgl. Seleskovitch, 1978: 31; AIIC, 2015), könnte dies zu einer möglicherweise unübersichtlicheren Darstellung des Transkripts führen, die auf die äußeren Umstände einer schnellen Vortragsweise und nicht auf einen Programmfehler hindeuten würde.

Der Fließtext wird jedenfalls auf diese Weise aufgeteilt und es sollte somit ein besserer Überblick geschaffen werden, der bei der anschließenden Wiedergabe der computergestützt angefertigten Mitschrift entscheidend ist. Ünlü (2023: 71) unterstreicht die Wichtigkeit eines segmentierten Texts für eine Verwendung im konsekutiven Arbeitsmodus.

Wenn die Rede zu Ende ist, kann schließlich per Klick auf „Stop Recognition“ die Aufnahme durch die integrierte automatische Spracherkennung bei Sight-Terp beendet werden. Im Anschluss erscheint auch die restliche bzw. vollständige maschinelle Übersetzung in die Zielsprache.

Was das Layout betrifft, so übernimmt die maschinelle Übersetzung auf der rechten Seite auch die Nummerierung, die sich in der Transkription befindet, nicht jedoch die gelben Markierungen der Named Entities bzw. Zahlen. In jedem Fall wird dadurch größtenteils gewährleistet, dass die Segmente der Ausgangs- und der Zielsprache auf gleicher Ebene angeordnet werden und sich somit parallel gegenüberstehen, um die Suche nach einzelnen Entsprechungen bzw. das Wiederfinden im Text zu erleichtern.

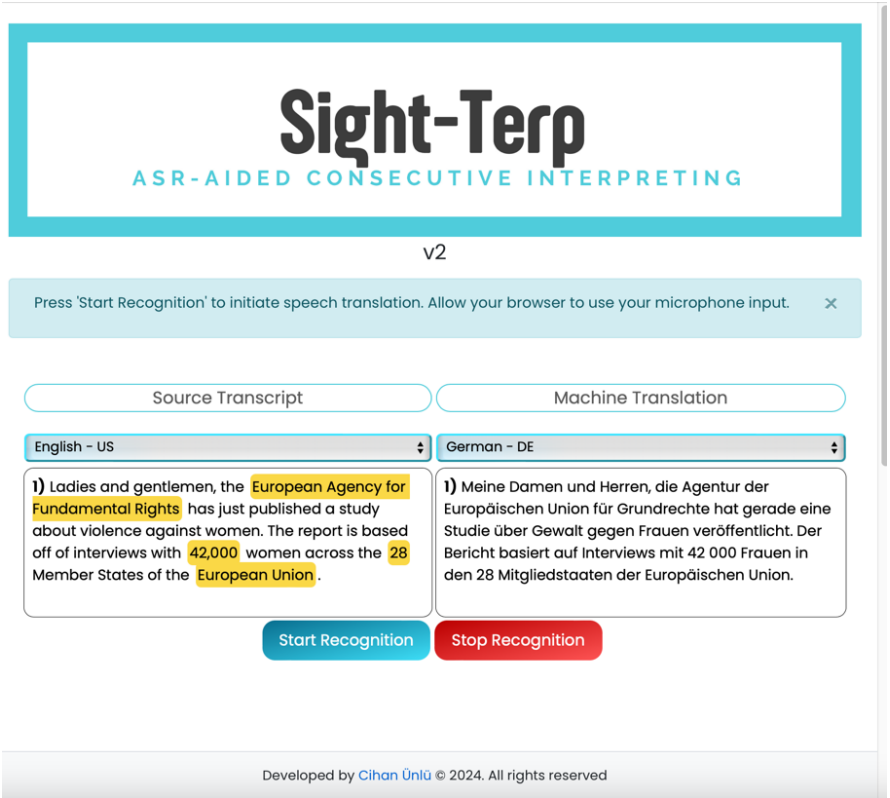


Abbildung 5: Benutzeroberfläche von Sight-Terp mit einem konkreten Beispiel (www.sightterp.net, Datum des Zugriffs: 10.06.2024)

Sight-Terp kann dabei sowohl am Laptop bzw. Computer als auch am Tablet-PC verwendet werden, solange der Bildschirm ausreichend groß ist, um ein einwandfreies Erkennen der Transkription und der parallel danebenstehenden Übersetzung zu ermöglichen.

Für das Erkennen von Videoaufnahmen ist dabei ein Zweitgerät notwendig, von dem die Rede abgespielt wird, damit die automatische Spracherkennung von Sight-Terp diese als sprachlichen Input richtig erfassen kann. In der vorliegenden Arbeit wird Sight-Terp auf einem Tablet-PC geöffnet, während auf einem Laptop die Ausgangsreden als Videos abgespielt wurden. Dadurch kann das Tablet auch besser angegriffen bzw. in die Hand genommen werden, was möglicherweise einem realen Einsatz näherkommt (vgl. Kapitel 4).

Eine weitere Funktionalität von Sight-Terp ist der dank Myscript integrierte digitale Notizblock (‚digital notepad'), welcher auf der Website direkt unter den beiden Spalten für die Transkription und die maschinelle Übersetzung zu finden ist. Auf einem Tablet-PC mit einem dafür vorgesehenen Stift lässt diese Funktionalität digitales Notieren zu (Ünlü, 2023: 73ff.). Da dieses Feature weder im empirischen Teil der Masterarbeit von Ünlü (2023) noch in der vorliegenden Arbeit Anwendung findet, soll darauf nicht genauer eingegangen werden.

Zudem sei angemerkt, dass Ünlü (2023: 75) selbst erkennt, dass die automatische Spracherkennung nicht „unfehlbar" ist, und Ünlü rät daher auch weiterhin zur gleichzeitigen Anfertigung von Notizen bei der Verwendung von Sight-Terp. Dies kann entweder durch den integrierten digitalen Notizblock erfolgen oder mit einem klassischen Notizblock und Stift.

3.3.3 Forschungsstand

In der Studie von Hamidi & Pöchhacker (2007) wurde die konsekutive Dolmetschleistung von professionellen Dolmetscher*innen im Rahmen eines Experiments im Modus SimConsec mit zwei Runden, einmal ohne und einmal mit digitalem Aufnahmegerät, analysiert. Es wurde herausgefunden, dass das computergestützte Konsekutivdolmetschen die Leistung unter diversen Gesichtspunkten, wie jenen der Flüssigkeit und der Entsprechung mit dem Ausgangstext, verbessern konnte.

In einer weiteren Studie erforschten Wang & Wang (2019), inwiefern die Dolmetschleistung, gemessen anhand der Genauigkeit und der Flüssigkeit, im computergestützten konsekutiven Arbeitsmodus beeinflusst wird. Dabei simulierten sie eine Anwendung für diesen Modus, indem eine Software zur Spracherkennung sowie maschinelle Übersetzung manuell kombiniert eingesetzt wurden (Wang & Wang, 2019: 123). Sie stellten fest, dass die Genauigkeit der Verdolmetschung verbessert werden konnte, dass aber über die Flüssigkeit keine Aussagen gemacht werden konnten, und sprechen sich für eine Weiterentwicklung von

Anwendungen für das computergestützte Konsekutivdolmetschen sowie eine Integration dieser Technologien in die Lehre aus (Wang & Wang, 2019: 135f.).

Chen & Kruger (2022: 400) überprüften in ihrer empirischen Studie die Effektivität des Einsatzes von automatischer Spracherkennung und maschineller Übersetzung (vgl. Abschnitt 3.1). Es wurden dabei Daten auf Prozess- und Produktebene im Sprachenpaar Chinesisch und Englisch gesammelt, welche 48 Verdolmetschungen umfassten, die mit einem Fokus auf den Aspekt der Genauigkeit hin analysiert wurden (Chen & Kruger, 2022: 409f.). Die Teilnehmenden verdolmetschten dabei in und aus ihrer Erstsprache, indem die im neu entwickelten Arbeitsmodus „CACI" vereinfacht ausgedrückten Schritte des Zuhörens und Nachsprechens und der anschließenden Wiedergabe durchgeführt wurden (vgl. Abschnitt 1.2.2).

Dabei fanden Chen & Kruger heraus, dass computergestütztes Konsekutivdolmetschen (kurz CACI) die kognitive Belastung aufgrund des Vorhandenseins einer maschinell erzeugten Übersetzung des Gesagten reduziert und die Qualität der Verdolmetschung in Hinblick auf die Genauigkeit, aber auch die Flüssigkeit im Vergleich zum klassischen Konsekutivdolmetschen erhöht (2022: 413f.). Chen & Kruger (2022: 414f.) erwähnen schließlich, dass die Sprachrichtung von der Erstsprache in die Fremdsprache die kognitive Belastung stärker reduziert als umgekehrt sowie dass die gewonnenen Erkenntnisse im Lichte einer erforderlichen besseren Vertrautheit mit dieser Technologie zu sehen sind.

Sight-Terp (Ünlü, 2023) liefert in diesem Zusammenhang einen sehr aktuellen Beitrag zum computergestützten Konsekutivdolmetschen. Ünlü (2023: 83ff.) wählte für die mit Sight-Terp durchgeführte Studie zunächst vier englische Ausgangsreden und überprüfte diese mittels Lesbarkeitsindizes auf Vergleichbarkeit. Nach Auswahl des Redematerials wurde dieses von einem englischen Muttersprachler aufgenommen und anschließend in Sinneinheiten aufgeteilt.

Als ersten Schritt führte Ünlü im Zuge des Experiments einen Pre-Test durch, bei dem die Teilnehmenden die Rede A1 über Erdbeben in Japan mit der klassischen Notizentechnik konsekutiv aus dem Englischen ins Türkische verdolmetschen sollten. Danach wurden die Teilnehmenden in die Verwendung von Sight-Terp eingeschult und das Tool wurde zu Übungszwecken bereitgestellt. Den Teilnehmenden wurden dabei die Reden auf einem Computer abgespielt und Sight-Terp wurde auf einem Tablet-PC abgespielt. Die Tatsache, dass Sight-Terp per Tablet eingesetzt wird, erlaubt auch eine flexible Handhabung des Geräts, da die Teilnehmenden beispielsweise per Berührung scrollen konnten und das Tablet, ähnlich wie einen Notizblock, in der Hand halten konnten. Neben den technischen Hilfsmitteln wurden zusätzlich ein Notizblock aus Papier und ein Stift zur Verfügung gestellt.

Im Zuge des anschließenden Post-Tests mussten die Teilnehmenden die Rede A2 über Erdbeben in Japan diesmal mit Sight-Terp verdolmetschen.

Diese Vorgehensweise wurde für die Reden B1 und B2 über Gewalt gegen Frauen wiederholt. Abschließend wurde das gesammelte Material anhand der Segmentierung in und Analyse von Sinneinheiten auf Genauigkeit überprüft. Die Flüssigkeit wurde ebenfalls u. a. durch die Anzahl der Pausen und Füllwörter gemessen. Schließlich führte Ünlü eine Umfrage durch, um die Eindrücke der Teilnehmenden zu dem Tool zu erheben.

Die Ergebnisse zeigen, dass zwar die Genauigkeit der Verdolmetschungen aller Teilnehmenden unter Zuhilfenahme von Sight-Terp gesteigert werden konnte, dass damit aber eine längere Dauer der Verdolmetschung und eine weniger flüssige Darbietung einherging (Ünlü, 2023: 93ff.).

Das folgende Kapitel beschäftigt sich mit der Methodik der vorliegenden Arbeit. Diese umfasst im Sinne der Replikation der Untersuchung von Ünlü (2023) ebenfalls ein Experiment sowie auch ein Leitfaden-Interview und wird im Folgenden näher vorgestellt.

4. Methodik

Im Folgenden wird die empirische Datenerhebung genauer beleuchtet, welche aus zwei Teilen besteht, nämlich einem Experiment, das als Replikation der Untersuchung von Ünlü (2023) angelegt ist, und einem anschließenden Leitfaden-Interview, um sowohl quantitative als auch qualitative Daten im Sinne des Mixed-Methods-Ansatzes (vgl. Hale & Napier, 2014: 210f.) zu sammeln. Chen & Kruger (2022: 400) behaupten in diesem Zusammenhang, dass es bislang zu wenig empirische Daten betreffend die Effektivität von Technologien im Dolmetschprozess gibt, um diese erfolgreich in der Dolmetschpraxis einzuführen. Die vorliegende Masterarbeit kann somit einen Beitrag zur weiteren Erforschung der Effektivität von Dolmetschtechnologien leisten. Die dafür vorgesehenen Forschungsfragen lauten:

1) Inwiefern kann die Verwendung von Sight-Terp beim Konsekutivdolmetschen im Sprachenpaar Englisch-Deutsch im Vergleich zur ausschließlichen Verwendung der klassischen Notizentechnik zu einer Verbesserung der Qualität führen, gemessen anhand der Genauigkeit der Wiedergabe von Sinneinheiten sowie der Verdolmetschung ausgewählter Problemauslöser?

2) Wie arbeiten Dolmetschstudierende insbesondere nach eigener Aussage mit Sight-Terp und wie schätzen sie ihre eigene konsekutive Dolmetschleistung mit bzw. ohne Sight-Terp ein?
    a. Inwiefern ändert sich die Dauer der Verdolmetschung mit Sight-Terp?
    b. Wie fertigen die Teilnehmenden ihre Notizen bei der gleichzeitigen Verwendung von Sight-Terp an?

Die zugehörige Forschungshypothese lautet:

1) Die Anwendung Sight-Terp verbessert die Qualität beim konsekutiven Dolmetschen im Sprachenpaar Englisch-Deutsch, gemessen anhand der Anzahl der fehlerfreien Sinneinheiten, sowie der Anzahl der korrekt gedolmetschten Problemauslöser im Vergleich zur Verdolmetschung ohne Sight-Terp.

Es besteht in diesem Sinne insbesondere Forschungsbedarf mit und an Studierenden, da universitäre Lehrveranstaltungen laut Fantinuoli (2018a: 163) das computergestützte

Dolmetschen nicht ausreichend abdecken. Die Teilnehmenden an diesem Experiment bilden daher fortgeschrittene Dolmetschstudierende am Zentrum für Translationswissenschaft (vgl. Abschnitt 4.1), womit durch die Teilnahme der Studierenden nicht nur relevante Daten für das bessere Verstehen des computergestützten Konsekutivdolmetschens gesammelt, sondern den Studierenden auch eine sehr aktuelle und möglicherweise neue Art des Dolmetschens anhand eines erst vor kurzem entwickelten Tools nähergebracht werden kann. Das Forschungsdesign soll daher kurz umrissen werden.

Der vorliegende empirische Teil orientiert sich in seiner Konzeption und Umsetzung an der produktorientierten Forschung nach Fantinuoli (2018a: 167f.), da sie im empirischen Teil anhand vordefinierter Parameter mögliche Einflüsse der CAI-Anwendung Sight-Terp auf die gedolmetschten Texte untersucht (vgl. Ünlü, 2023: 6). Der komparativen Komponente wird durch zwei Durchgänge, jeweils abwechselnd mit und ohne Verwendung der Softwareanwendung Sight-Terp, Rechnung getragen.

Die der Datenerhebung folgende Beurteilung der Verdolmetschungen erfolgt einerseits quantitativ mittels der erweiterten Fehleranalyse nach Zhao (2021) anhand eines in Sinneinheiten aufgeteilten schriftlichen Textmaterials der transkribierten Originalreden und Verdolmetschungen, wobei ein spezieller Fokus den Problemauslösern gilt, sowie andererseits qualitativ mittels eines anschließenden Leitfadeninterviews, bei dem versucht wird, prozessorientierte Daten bzw. Informationen vom tatsächlichen Arbeiten mit der Anwendung von den Teilnehmenden zu erhalten.

Diese Art der Forschung, bei der nachträglich ein Rekonstruieren der ausgeführten Prozesse und der dabei entstandenen Überlegungen erfragt wird, ist im Bereich der intro- bzw. retrospektiven Forschung angesiedelt (Vik-Tuovinen, 2002: 63). Dabei merkt Vik-Tuovinen (2022: 63) in Hinblick auf das Simultandolmetschen an, dass keine gleichzeitige Selbstbeobachtung bzw. Introspektion während der Verdolmetschung möglich sei und auch nicht alle Entscheidungen so bewusst getroffen werden, dass man sich im Nachhinein an sie erinnern könnte. Genauso ist es während des Konsekutivdolmetschens nicht möglich, laut auszusprechen, was zuerst und danach gemacht wird. Trotz dieses Defizits soll die nachträgliche Erfragung von Abläufen während der Verdolmetschung mit Sight-Terp im Rahmen des Leitfaden-Interviews einen Einblick in (bewusste) Arbeitsprozesse sowie weitere diesbezügliche Überlegungen geben.

4.1 Teilnehmende

Insgesamt nahmen 10 Personen an dem Experiment teil. Alle Teilnehmenden gaben an, fortgeschrittene Dolmetschstudierende ab dem 4. Semester am Zentrum für Translationswissenschaft zu sein, und wiesen in ihrer Sprachenkombination neben weiteren Sprachen sowohl Englisch als auch Deutsch auf, wenn auch in unterschiedlicher Reihenfolge nach der A/B/C-Kategorisierung in aktive bzw. passive Sprachen.

Den Teilnehmenden wurde aus datenschutzrechtlichen Gründen eine anonymisierte Bezeichnung, abgekürzt als TN1, TN2, etc. für Teilnehmende eins, zwei, etc. im Rahmen der Datenanalyse vergeben.

Das Experiment zeichnet sich durch ein sich wiederholendes Messverfahren aus (‚repeated measures design'), das bedeutet, dass die gleichen Teilnehmenden (vgl. ‚within-subject design') beiden gemessenen Bedingungen abwechselnd in jeweils zwei Durchgängen ausgesetzt werden (Mellinger & Hanson, 2016: 7). Dadurch können individuelle Unterschiede zwischen den Teilnehmenden, wie unterschiedliches Vorwissen, neutralisiert werden.

In der folgenden Zusammenstellung soll ein Überblick über das Profil sowie die genauen Tage und die Uhrzeiten der jeweiligen Experimente gegeben werden. Zur besseren Übersichtlichkeit werden diese in Tabelle 1 veranschaulicht. Weitere angeführte Sprachen, die sich in den Sprachenkombinationen der Teilnehmenden finden, sind Deutsch (DE), Englisch (EN), Italienisch (IT), Spanisch (ESP), Russisch (RUS), Bosnisch-Kroatisch-Serbisch (BKS), Französisch (FR). Dabei zeigt sich außerdem, dass alle Teilnehmenden zum Zeitpunkt der Durchführung fortgeschrittene Dolmetschstudierende ab dem 4. Semester waren und somit allesamt aufgrund zahlreicher belegter Lehrveranstaltungen wichtige Erfahrungen im konsekutiven Dolmetschen zum Experiment mitbrachten.

Tabelle 1: Profilüberblick TN1-TN10

TN-ID	Semester	Schwer-punkt	Sprachen-kombination	Datum und Uhrzeit des Experiments
TN1	4	Konferenz-dolmetschen	IT (A), DE (B), EN (C)	04.05.2024, von 10.30-11.30 Uhr
TN2	5	Konferenz-dolmetschen	DE (A), IT (B), EN (C)	06.05.2024, von 15.00-16.00 Uhr

TN3	4	Konferenz-dolmetschen	DE (A), EN (B), FR (C)	07.05.2024, von 14.30-15.30 Uhr
TN4	5	Dialog-dolmetschen	DE (A), FR (B), EN (B)	09.05.2024, von 14.00-15.00 Uhr
TN5	4	Konferenz-dolmetschen	DE (A), EN (B), ESP (C)	10.05.2024, von 8:30-9.30 Uhr
TN6	4	Konferenz-dolmetschen	DE (A), ESP (B), EN (C)	10.05.2024, von 11.30-12.30 Uhr
TN7	6	Konferenz-dolmetschen	RUS (A), DE (B), EN (C)	10.05.2024, von 15.30-16.30 Uhr
TN8	6	Dialog-dolmetschen	DE (A), FR (B), EN (B)	11.05.2024, von 10.00-11.00 Uhr
TN9	5	Konferenz-dolmetschen	DE (A), EN (B), FR (C), ESP (C)	13.05.2024, von 13.00-14.00 Uhr
TN10	5	Konferenz-dolmetschen	DE (A), BKS (B), EN (C)	21.05.2024, von 16.45-17.45 Uhr

4.2 Experiment

Da das in der vorliegenden Arbeit durchgeführte Experiment eine Replikation jenes von Ünlü (2023) (vgl. Abschnitt 3.3.3) unter ähnlichen Bedingungen darstellt, soll im Folgenden auf die Unterschiede zwischen den beiden Experimenten eingegangen werden.

Zunächst war das analysierte Sprachenpaar in der vorliegenden Arbeit Englisch-Deutsch. Außerdem wurde bereits eine Woche vor dem Experiment das Tool den Teilnehmenden präsentiert und gemeinsam ausprobiert, damit neben der Beantwortung noch auftauchender Fragen die Teilnehmenden vor allem die Gelegenheit hatten, sich selbständig zuhause damit vertraut zu machen bzw. weitere Fragen vorab zu klären.

Am Tag des Experiments wurde ein Pre-Test zu einem nicht mit den eigentlichen Testreden zusammenhängenden Thema durchgeführt, da dieser nur dem Aufwärmen für die Verwendung des Tools diente und inhaltliche Einflüsse auf die Verdolmetschung während des Experiments möglichst ausgeschlossen werden sollten.

Insgesamt erhielten alle Teilnehmenden jeweils nur zwei Reden zu zwei Themen und diese wurden abwechselnd mit bzw. ohne Sight-Terp konsekutiv verdolmetscht. Bei Ünlü

waren es pro Durchgang aufgrund der beiden Pre-Tests in Summe vier zu dolmetschende Reden und für alle Teilnehmenden waren es bei Ünlü auch immer die gleichen Reden, die mit bzw. ohne technologische Unterstützung durch diese verdolmetscht wurden.

In Hinblick auf die darauffolgende qualitative Datenerhebung wurde bei Ünlü ein Fragebogen erstellt und an die Teilnehmenden verteilt; in dem vorliegenden Experiment wurde ein Leitfaden-Interview mit Fragen zur offenen Diskussion durchgeführt.

Sowohl die jeweiligen Verdolmetschungen der Teilnehmenden sowie auch die im Anschluss durchgeführten Leitfaden-Interviews wurden mithilfe des Textverarbeitungsprogramms Word zunächst mit der automatischen Diktierfunktion automatisch transkribiert und dann bei einem nochmaligen Abspielen der Aufnahmen überarbeitet. Da prosodische Merkmale, wie beispielsweise die Intonation und Häsitationslaute sowie Redundanzen, lautes Sprechen etc. nicht Untersuchungsgegenstand waren, wurde die Transkription nicht aufgrund einer dahingehenden Methode zur Analyse dieser Phänomene durchgeführt, es wurden vielmehr die für diese Zwecke relevanten Transkriptionsregeln von Kuckartz & Rädiker (2022) herangezogen (vgl. Abschnitt 4.3.3).

Zusammenfassend kann folgender Überblick über das Forschungsdesign gegeben werden: Die unabhängige Variable stellt das (Nicht-)Vorhandensein der Anwendung Sight-Terp und die untersuchte abhängige Variable die Qualität dar, definiert und analysiert anhand der Bestimmung von Sinneinheiten mit speziellem Fokus auf ausgewählte Problemauslöser. Unberücksichtigt blieben weitere Aspekte, wie z. B. die Flüssigkeit der Wiedergabe.

Die Arbeitsbedingungen für die Teilnehmenden können als der konsekutive Arbeitsmodus und die Sprachrichtung vom Englischen ins Deutsche beschrieben werden. Mögliche weitere unabhängige Parameter, die einen Einfluss auf die Verdolmetschung haben könnten, wären die mehr oder weniger einwöchige Erfahrung mit dem Tool Sight-Terp sowie die Erfahrung beim Dolmetschen in der Sprachenkombination Englisch-Deutsch und das Vorwissen zu den beiden Themen des Experiments. Da den Teilnehmenden jeweils eine Woche vor dem Experiment sowohl das Tool als auch die Themen präsentiert wurden, hatten alle die gleichen Ausgangsbedingungen für die Vorbereitung, und da alle Dolmetschstudierende zum Zeitpunkt der Durchführung mindestens im 4. Semester waren, kann davon ausgegangen werden, dass trotz eines unterschiedlichen Sprachenkanons (Kategorisierung in A/B/C-Sprache) doch ein vergleichbares Maß an Dolmetscherfahrung vorhanden war. Somit wurde auch versucht, den Einfluss nicht untersuchter unabhängiger Parameter im Rahmen des Experiments so gering wie möglich zu halten, um das eigentliche Forschungsinteresse in den Mittelpunkt der Untersuchung zu rücken.

4.2.1 Redematerial

Das Redematerial (vgl. Anhang I) entstammt dem replizierten Experiment von Ünlü (2023), wobei in einem ersten Schritt alle ursprünglich vier englischen Ausgangsreden von einem amerikanischen Muttersprachler als Video- und Audioaufnahmen aufgezeichnet wurden. Dazu wurde der Redner darum gebeten, die schriftlich vorhandenen Texte so frei wie möglich zu präsentieren, das heißt mit entsprechender Intonation am Satzende und am Beginn eines neuen Satzes bzw. beim Hervorheben von Informationen. Zudem wurde auf eine möglichst klare Aussprache Wert gelegt, nicht zuletzt in Hinblick auf die Geschwindigkeit.

Am Anfang jeder Rede wurden die Nummer und der jeweilige Titel zum besseren Einordnen und letztlich das Signal „I will start the speech now" gegeben, damit die Dolmetscher*innen wussten, ab wann sie ihre Verdolmetschung beginnen sollten. Dieser Teil wurde dementsprechend nicht gedolmetscht. Da die tatsächlich ausgesprochenen Wörter der Reden zudem minimal vom ursprünglichen Redematerial abwichen, wurden im Anschluss alle Reden auf den tatsächlich Wortlaut hin analysiert und erneut transkribiert.

Danach wurden die Reden miteinander verglichen, nicht zuletzt in Hinblick auf die Vortragsgeschwindigkeit und -weise, wie etwa die Prosodie bzw. die Intonation. Alle vier Reden wiesen eine Geschwindigkeit von 150-160 Wörtern pro Minute auf sowie eine Dauer von ungefähr 3 Minuten, genauer gesagt waren es bei der in einem späteren Schritt ausgewählten Rede B1 exakt 3:20 Minuten und bei der Rede A2 genau 3:10 Minuten.

Dabei wurde wie schon erwähnt festgestellt, dass die Rede B1 zum Thema Gewalt gegen Frauen mit einer Wortanzahl von 510 Wörtern und die Rede A2 zum Thema Erdbeben in Japan mit einer Wortanzahl von 452 Wörtern unter den Gesichtspunkten der Prosodie bzw. Intonation am vergleichbarsten waren. Sie wurden daher für das Experiment ausgewählt. Es wurde darauf geachtet, dass die beiden Reden für das Experiment in jedem Durchgang jeweils ein anderes Thema betreffen, damit kein möglicher Übungseffekt von der ersten Runde als weitere unabhängige Variable die Ergebnisse der Verdolmetschung der zweiten Runde beeinflussen konnte.

Für die Reden B1 und A2 wurden anschließend mithilfe eines webbasierten Gunning Fog Index[1] die Schwierigkeitsindizes berechnet. Es wurde der Gunning Fog Index exemplarisch ausgewählt, da dieser einen der am weitesten verbreiteten Indizes zur Einschätzung der Lesbarkeit bzw. Unlesbarkeit von Texten darstellt (vgl. Świeczkowski & Kułacz, 2021: 627).

[1] abrufbar unter http://gunning-fog-index.com/

Beim Gunning Fog Index wird der Schwierigkeitsgrad insbesondere anhand der Anzahl der Wörter und, genauer, der mehr als dreisilbigen Wörter kalkuliert, wobei ein höherer Wert einen schwierigeren Text impliziert. Diese Art der Berechnung ist auch für die Beurteilung der Schwierigkeit beim Dolmetschen interessant, da eine durch längere Wörter erschwerte Lesbarkeit auch den Vortrag eines somit schwierigeren schriftlichen Ausgangstexts beeinflusst und folglich die erste Phase des kognitiven Prozesses beim Konsekutivdolmetschen (vgl. Gile, 2009: 175f.), wie etwa das Zuhören, Merken und Notieren längerer, dadurch möglicherweise komplexerer Wörter beeinträchtigen könnte. Die Berechnung ergab vergleichbar hohe Werte von 11,55 für die Rede B1 und 13,09 für die Rede A2. Ünlü (2023: 79f.) kam bei der Berechnung des Gunning Fog Index bei der Rede B1 auf einen Wert von 11,24 und bei der Rede A2 auf einen Wert von 11,40. Unterschiede im Ergebnis sind evtl. auf leicht unterschiedliche Berechnungsmethoden zurückzuführen.

Zudem wurden die Audioaufnahmen zunächst mittels Audacity[1] auf längere Pausen, ab einer bis maximal zwei Sekunden, hin überprüft (vgl. Abbildung 6). Dies konnte dank einer Vergrößerung der Ansicht, die eine Unterteilung der Segmente in Sekundenabschnitte zuließ, visuell erfolgen. Dabei wurde festgestellt, dass keine der beiden Aufnahmen Pausen enthält, die länger als 2 Sekunden dauern.

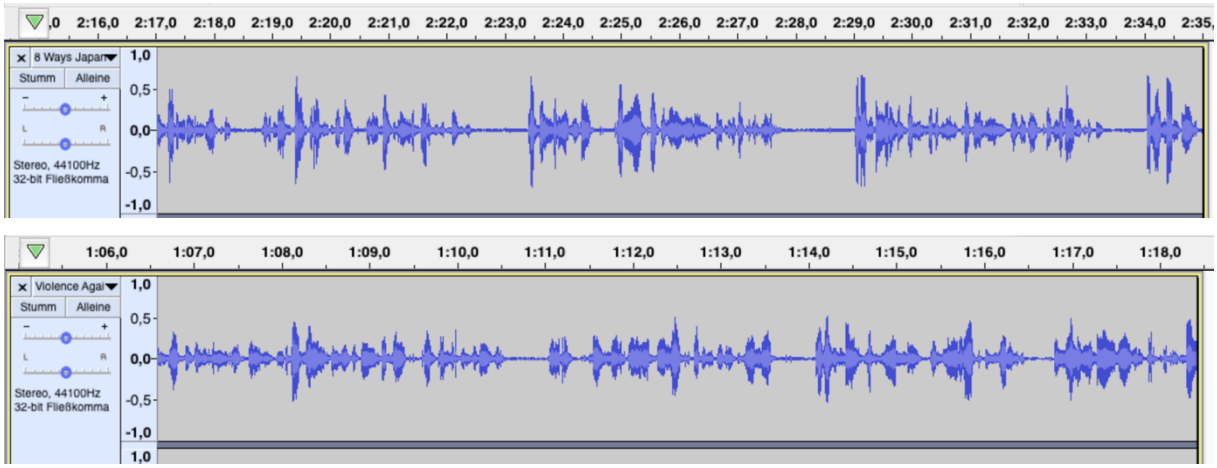


Abbildung 6: Bildschirmaufnahmen der Reden in Audacity

In einem nächsten Schritt wurden als übergreifender Geschwindigkeitsindikator die Wörter pro Minute für beide Reden berechnet. Für die Rede B1 ergab diese Methode einen Wert von 160 Wörtern pro Minute, für die Rede A2 waren dies 150 Wörter pro Minute. Diese Werte

[1] online frei verfügbar unter https://www.audacityteam.org

ähneln auch den von Seleskovitch (1978: 31) angesprochenen 150 Wörtern pro Minute, sind jedoch ein Zeichen für eine sehr rasche Vortragsgeschwindigkeit der hier vorgetragenen Redebeiträge. Die Geschwindigkeit wäre somit als ein potenzieller Problemauslöser für die Teilnehmenden einzustufen. Gleichzeitig könnte dieser Problemauslöser dazu geführt haben, dass die Teilnehmenden bei der Verdolmetschung im Durchgang mit Sight-Terp tatsächlich davon Gebrauch machten, was wiederum dem Untersuchungsgegenstand dienlich war.

Zusätzlich wurden die beiden ausgewählten Reden im Vorhinein, d. h. vor der Durchführung des Experiments, abgespielt und von Sight-Terp transkribiert und übersetzt. Die Bildschirmaufnahmen der ASR-Transkriptionen und maschinellen Übersetzungen spiegeln aufgrund der Tatsache, dass das Redematerial immer dieselben Videoaufnahmen sind, die tatsächlichen Bedingungen aller Teilnehmenden wider. Somit konnte durch eine Gegenüberstellung der ausgangssprachlichen Transkription mit dem ursprünglichen Originaltext sowie durch die genaue Durchsicht der deutschen Übersetzung eine Fehlerrate (‚word error rate', kurz WER) des Tools Sight-Terp für die ausgewählten Reden in der Sprachenkombination Englisch-Deutsch ermittelt werden.

Die Resultate der Berechnung der WER für die mittels ASR automatisch generierten englischen Transkripte sind folgende: 1,18 % für die Rede B1 und 0,88 % für die Rede A2, wobei hier tatsächliche Transkriptionsfehler (z. B. statt ‚I thought' wurde ‚without' transkribiert) gezählt wurden. Trotz einer raschen Vortragsweise könnte dieser geringe Wert auf die Tatsache zurückzuführen sein, dass die Aufnahmen von einem amerikanischen Muttersprachler angefertigt und sehr klar und deutlich artikuliert wurden. Ob die automatische Spracherkennung eine ebenso geringe Fehlerrate unter anderen Bedingungen, wie etwa bei Akzenten von ELF-Sprecher*innen, aufweisen würde, wäre an dieser Stelle eine weitere interessante Fragestellung, auch um der Forderung nachzukommen, dass automatische Spracherkennung sprecherunabhängig sein sollte (vgl. Fantinuoli 2017: 2).

Die Berechnung der WER wurde dann ebenso für die maschinelle Übersetzung durchgeführt, indem die automatische Transkription sowie ihre maschinelle Übersetzung dank einer Bildschirmaufnahme von Sight-Terp für die jeweilige Rede mit der Originalrede verglichen wurde. Eine Ähnlichkeit der Fehlerwerte für die maschinelle Übersetzung bestand insofern, dass diese unmittelbar von der automatischen Spracherkennung abhing und somit Fehler in der deutschen Version zumeist auf eine fehlerhafte englische Transkription zurückzuführen waren. Jedoch gab es auch weitere Problemstellen, wie Übernahmen von Elementen des englischen Originals in die deutsche Übersetzung (z. B. ‚Risikoexposition' für ‚risk exposure' oder ‚Züge einfrieren' für ‚freeze trains', sowie syntaktische Schwierigkeiten wie ‚also zum

Schluss ich würde sagen‘ etc.), sowie Ausdrücke, die womöglich nicht ohne eine vorherige mündliche ad hoc Bearbeitung für eine Konsekutivdolmetschung herangezogen werden könnten, da sie sich zu sehr am Ausgangstext orientieren und wenig idiomatisch sind.

Die Frage, inwiefern komplette Übernahmen (wie Backlash im Deutschen) als Fehler gewertet werden sollten, hängt andererseits stark vom Zielpublikum der Verdolmetschung ab und sollte für die jeweilige Situation individuell geklärt werden (vgl. Abschnitt 4.2.3).

Neben den oben genannten Fehlern gab es auch die Problematik, dass beispielsweise Sätze an Stellen aufgeteilt oder zusammengeschrieben wurden, wo dies aufgrund der logischen Abfolge im Text nicht passend bzw. für eine anschließende Verdolmetschung nicht übersichtlich war. Ebenso wurden vereinzelt keine Markierungen von Zahlen in der englischen Transkription vorgenommen, die dann natürlich nur im englischen Transkript als Fehler gewertet wurden, da sie für die maschinelle Übersetzung generell nicht vorgesehen sind. Diese das Layout betreffenden Fehler wurden separat berechnet.

Die Aufteilung des Transkripts in nummerierte Absätze war möglicherweise aufgrund des schnellen Vortragstempos beeinträchtigt, da viele Abschnitte, v. a. bei der Aufzählung der 8 angesprochenen Punkte in der Rede A2 über Erdbeben in Japan zusammengezogen waren, die eigentlich einen neuen Absatz hätten bilden sollen. Da Sight-Terp in diesem Fall jedoch die Punkte Erstens, Zweitens, etc. in der ausgangssprachlichen Transkription farblich markierte, war dadurch trotz einer falschen Absatzgliederung zumindest im englischen Transkript ein gewisser Überblick gegeben.

Die Tabellen 2 und 3 stellen abschließend einen Überblick der analysierten Parameter sowie die genauen Verdolmetschungsdurchgänge der Teilnehmenden dar und dienen der besseren Vergleichbarkeit.

Tabelle 2: Überblick über die Eigenschaften der Reden

	Rede B1: Gewalt gegen Frauen	Rede A2: Erdbeben in Japan
Wortanzahl	510 Wörter	452 Wörter
Dauer in Minuten (min.)	3:20 min.	3:10 min.
Gunning Fog Index	11,55	13,09
Geschwindigkeit in Wörter pro Minute (wpm)	160 wpm	150 wpm

Fehlerrate (WER) für die Transkription per ASR	1,18 %	0,88 %
Fehlerrate (WER) für die maschinelle Übersetzung	3,33 %	1,99 %
Layoutfehlerrate für die Transkription per ASR	0,59 %	1,33 %
Markierungsanzahl für die Transkription per ASR	16	17
Layoutfehlerrate für die maschinelle Übersetzung	0,39 %	0,44 %
Anzahl der Absätze für die Transkription per ASR und die maschinelle Übersetzung	8	8

Tabelle 3: Überblick über die Durchgänge mit/ohne Sight-Terp

	Rede A2: Erdbeben in Japan	Rede B1: Gewalt gegen Frauen
TN1	mit Sight-Terp	ohne Sight-Terp
TN2	ohne Sight-Terp	mit Sight-Terp
TN3	mit Sight-Terp	ohne Sight-Terp
TN4	ohne Sight-Terp	mit Sight-Terp
TN5	mit Sight-Terp	ohne Sight-Terp
TN6	ohne Sight-Terp	mit Sight-Terp
TN7	mit Sight-Terp	ohne Sight-Terp
TN8	ohne Sight-Terp	mit Sight-Terp
TN9	mit Sight-Terp	ohne Sight-Terp
TN10	ohne Sight-Terp	mit Sight-Terp

4.2.2 Vorgehensweise

Zunächst wurde mit den Teilnehmenden zirka eine Woche vor dem geplanten Termin des Experiments per Zoom ein Online-Treffen vereinbart, bei dem der Ablauf des Experiments sowie Sight-Terp genau vorgestellt und gemeinsam getestet wurden. In diesem Rahmen konnten bereits erste Versuche mit dem Tool gemacht sowie offene Fragen gestellt werden. Dazu registrierten sich die Teilnehmenden kostenlos mit ihrer E-Mail-Adresse auf der Webseite[1] und konnten bereits kurz darauf die Ausgangs- und Zielsprache auf der Benutzeroberfläche auswählen (vgl. Abschnitt 3.3.2).

Als Testrede wurde den Teilnehmenden eine YouTube-Rede zum Thema Menschenrechte[2] auf ein Zweitgerät geschickt, welche sie noch im Rahmen dieses Treffens abspielten und so die Funktionsweise von und das Arbeiten mit Sight-Terp austesten konnten. Außerdem bekamen sie so die Möglichkeit, bereits eine computergestützte konsekutive Verdolmetschung des geschickten Videos durchzuführen.

Nachdem einige Fragen zum Ablauf des Experiments und zur Anwendung geklärt wurden, erfuhren die Teilnehmenden die beiden Themen für die Reden beim Experiment, nämlich „Gewalt gegen Frauen“ und „Erdbeben in Japan“, und wurden gebeten, sich terminologisch, etwa mittels Glossarerstellung oder Symbolfindung, wie auch technisch vorzubereiten, d. h. sie sollten Sight-Terp bereits vor dem Experiment selbst ausprobieren, um etwas Erfahrung damit zu sammeln.

Den Teilnehmenden wurde auch gesagt, sie sollten für das Experiment etwas zum Notieren mitnehmen, einerseits für die Runde ohne Sight-Terp, andererseits wurde auch dazugesagt, dass selbst unter Verwendung von Sight-Terp daran gedacht werden soll, dass Sight-Terp Fehler in der Transkription machen kann (vgl. Ünlü, 2023: 75) und beim Vorhandensein von schriftlichem Material für die Verdolmetschung nach wie vor das tatsächlich Gesagte gilt (vgl. Pöchhacker, 2016: 20). Somit sollten sie sich im Zweifelsfall nicht „blind“ einzig und allein auf die Anwendung verlassen. Das bedeutet auch, dass im Zuge der computergestützten Verdolmetschung eigene Notizen angefertigt werden konnten.

Die Experimente fanden für die jeweiligen Teilnehmenden an unterschiedlichen Tagen im Mai 2024 und alle im Rahmen eines persönlichen Treffens statt. Der erste Termin war am 04.05.2024 und der letzte am 21.05.2024.

---

[1] abrufbar unter https://www.sightterp.net
[2] abrufbar unter https://www.youtube.com/watch?v=nDgIVseTkuE

Am jeweiligen Tag des Experiments wurden die Teilnehmenden zunächst gebeten, eine Einverständniserklärung für die Aufzeichnung und die anonymisierte Verwendung ihrer transkribierten Verdolmetschungen im Rahmen dieser Masterarbeit zu unterzeichnen. Dann wurde Sight-Terp auf einem Tablet-PC der Marke Apple geöffnet, die Sprachen English (US) und German (DE) wurden in einem Drop-Down-Menü auf der Benutzeroberfläche ausgewählt und das Tablet waagrecht und mittig aufgestellt, um eine gute Lesbarkeit sowie möglichst realitätsnahe Bedingungen zu gewährleisten.

Zunächst wurde zum Aufwärmen ein Pre-Test durchgeführt. Zu diesem Zwecke wurde eine YouTube-Rede zum Thema Europäische Union[1] gewählt und auf einem Laptop abgespielt, welcher rechts neben dem Tablet-PC aufgestellt wurde, um eine gute Sicht auf beide Bildschirme, nämlich Sight-Terp und das jeweils aufgenommene Video der Rede, zu ermöglichen. Die Pre-Test-Rede wurde gewählt, da sie ein im Rahmen des Studiums, insbesondere im Bereich des Konferenzdolmetschens, oft behandeltes Thema anspricht, welches auch einige, in diesem Kontext nicht unbekannte, Problemauslöser, wie Terminologie und Zahlen sowie Eigennamen enthält. Die Pre-Test-Rede dauerte 3:40 Minuten und wurde mit einer Geschwindigkeit von ca. 160 Wörtern pro Minute vorgetragen, sie war also ähnlich lang und schnell wie die Reden des Experiments. Diese erste Runde wurde nicht aufgenommen und ist kein Bestandteil der Untersuchung, sie diente einzig und allein der nochmaligen Auseinandersetzung mit dem Tool, dem Aufwärmen für das Dolmetschen sowie der Klärung von möglicherweise noch auftauchenden Fragen. Thematisch sollte diese Vorübung keinen Einfluss auf die Ergebnisse des darauffolgenden Experiments haben, daher fiel die Wahl wie oben beschrieben auf ein nicht mit den im Anschluss zu verdolmetschenden Videos zusammenhängendes und nicht vorher angekündigtes Thema.

Danach begann das eigentliche Experiment. Die Audioaufnahme wurde jetzt dank eines Diktiergeräts auf einem Smartphone gestartet. Die erste Runde im Rahmen des Experiments betraf für alle gleichermaßen die computergestützte konsekutive Verdolmetschung, das heißt die Teilnehmenden durften Sight-Terp, und ggf. ihre eigenen Notizen, einsetzen. Die dafür verwendete Rede war für TN1 die Rede A2 über Erdbeben in Japan, für TN2 war dies die Rede B1 über Gewalt gegen Frauen, usw. und die zweite Runde umfasste die konsekutive Verdolmetschung ohne Sight-Terp, wofür wiederum für TN1 die Rede B1 verwendet wurde und für TN2 die Rede A2, usw. Die ausgewählten Reden wurden also abwechselnd von den Teilnehmenden mit bzw. ohne Technologieeinsatz gedolmetscht, um mögliche Einflüsse auf

[1] abrufbar unter https://www.youtube.com/watch?v=SFeB9fMGJ9k

die Verdolmetschung, die auf unterschiedliche Schwierigkeitsgrade der Reden oder unterschiedliches Vorwissen zurückzuführen sind, möglichst auszugleichen.

Bei zwei Experimenten, nämlich bei TN7 und TN8, gab es technische Probleme, da nach dem Öffnen der Anwendung Sight-Terp per Klick auf „Start Recognition“ die automatische Spracherkennung nicht gestartet werden konnte. Bei TN7 tauchte dieses Problem erst nach dem Pre-Test mit Sight-Terp auf, der noch mit Sight-Terp durchgeführt werden konnte. Bei TN8 konnte auch der Pre-Test nicht durchgeführt werden. Stattdessen wurde die Anwendung in diesen beiden Fällen simuliert, das heißt es wurden die vorab zur Analyse gemachten Bildschirmaufnahmen der Transkription und der maschinellen Übersetzung für die beiden mit Sight-Terp zu verdolmetschenden Reden jeweils parallel in ein Word-Dokument kopiert und den betroffenen Teilnehmenden so für ihre konsekutive Verdolmetschung nach dem Abspielen der Reden zur Verfügung gestellt. Da es sich beim Ausgangsmaterial um aufgenommene Videos handelt, sind die Bildschirmaufnahmen der Transkriptionen und der maschinellen Übersetzung mit jenen im tatsächlichen Experiment ident.

Eine Kontaktaufnahme mit Ünlü am 12.05.2024 per E-Mail führte dankenswerterweise rasch zu einer Lösung des Problems, denn indem Sight-Terp im Incognito-Modus des Browsers aufgerufen wird, können die oben beschriebenen technischen Probleme umgangen werden. So wurde auch noch für die letzten beiden Experimente mit den Teilnehmenden TN9 und TN10 verfahren, die Sight-Terp im Zuge des Experiments wieder so wie ursprünglich vorgesehen verwenden konnten.

Nachdem beide Durchgänge beendet waren, wurde die Aufnahme nicht sofort angehalten, sondern es erfolgte ein nahtloser Übergang zum anschließenden Leitfaden-Interview, auf das im Abschnitt 4.3 genauer eingegangen wird.

4.2.3 Analysemethode

Für die Analyse der Ausgangstexte und ihrer Verdolmetschungen wurden diese zunächst in Sinneinheiten und Problemauslöser aufgeteilt und für die Anfertigung eines Analyse- und Bewertungsrasters für die konsekutiven Verdolmetschungen herangezogen, wobei dieser jeweils in ein Excel-Sheet pro Teilnehmenden eingefügt wurde.

Eine Sinneinheit sollte für die Unterteilung und die Bewertung in ihrer Hauptaussage in sich abgeschlossen sein, wurde aber, im Gegensatz zu einem Satzglied, nicht zwangsläufig auf ein einzelnes Subjekt, Verb, eine Konjunktion etc. heruntergebrochen. Jedoch kam es aufgrund einer detaillierten Unterteilung zu mehr Sinneinheiten als bei Ünlü (vgl. 2023: 80). Die Unterteilung erfolgte mithilfe von Fragewörtern bzw. W-Fragen (vgl. Gillies, 2017: 37f.).

Daneben wurden Problemauslöser absatzweise kenntlich gemacht, darunter insbesondere Zahlen. Dabei zeigte sich, dass die Rede A2 über Erdbeben in Japan einen leicht höheren Anteil an Problemauslösern enthält (vgl. Tabelle 4). Der folgende Auszug aus der Rede A2 über Erdbeben in Japan mit sieben Sinneinheiten und dem Sichtbarmachen der Problemauslöser sei als Beispiel angeführt.

In 1995,	*(SE: wann? / PT: Zahl)*
the city of Kobe was struck	*(SE: wer/ was? / PT: Eigenname*)
by a completely devastating earthquake,	*(SE: wovon genau?)*
which killed 5,000	*(SE: wie viel? / PT: Zahl)*
people	*(SE: wessen?)*
and destroyed tens of thousands	*(SE: wie viel? / PT: Zahl)*
of homes.	*(SE: wessen?)*

Auf diese Weise wurden beide durch die Teilnehmenden verdolmetschten Reden analysiert, da so neben den Problemauslösern auch stufenweise die in der Verdolmetschung wiedergegebenen Sinneinheiten beurteilt werden konnten.

Diese Vorgehensweise ergab nun für die Rede B1 über Gewalt gegen Frauen 159 Sinneinheiten und für die Rede A2 über Erdbeben in Japan 135 Sinneinheiten. Zum Vergleich betrugen die Werte bei Ünlü (2023: 80) je 159 und 127 Sinneinheiten.

Die folgende Zusammenstellung soll einen Überblick über die Ergebnisse der Analysen liefern und zeigt u. a., dass sich die Sinneinheiten und Problemauslöser in ihrem Anteil an der Gesamtwortanzahl der jeweiligen Rede mit 31,2 % und 9,4 % für die Rede B1 sowie 29,9 % und 12,6 % für die Rede A2 leicht unterscheiden.

Tabelle 4: Überblick über die Sinneinheiten und Problemauslöser in den Reden B1 und A2

	Rede B1: Gewalt gegen Frauen (510 Wörter)	Rede A2: Erdbeben in Japan (452 Wörter)
Anzahl der Sinneinheiten	159	135
Anteil der Sinneinheiten an der Gesamtwortanzahl	31,2 %	29,9 %

Anzahl der Problemauslöser (inkl. Named Entities)	48, davon: 14x Zahlen, 5x Aufzählungselemente, 12x Eigennamen, 17x einzelne Termini	57, davon: 19x Zahlen, 8x Aufzählungselemente, 19x Eigennamen, 11x einzelne Termini
Anteil der Problemauslöser (inkl. Named Entities) an der Gesamtwortanzahl	9,4 %	12,6 %

Im Folgenden wird der Aufbau des Bewertungs- und Analyserasters (vgl. Anhang II) anhand der einzelnen darin vorkommenden Spalten sowie die zur Beurteilung der Verdolmetschungen in dieser Spalte durchgeführte Vorgehensweise erläutert.

Spalte Sinneinheit

Zur Analyse der Ausgangstexte sowie zur anschließenden Evaluierung der Genauigkeit der Verdolmetschungen wurden zunächst die beiden englischen Ausgangsreden in Sinneinheiten gegliedert, da diese Vorgehensweise beim Anfertigen der Notizen für eine konsekutive Verdolmetschung in Hinblick auf das Notieren der gedanklichen Struktur (vgl. Abschnitt 1.2.2) zu bevorzugen ist und zudem als Bewertungsbasis für die Genauigkeit einzelner Segmente eine Gliederung in diese sinntragenden Abschnitte ermöglichte.

Bei der Unterteilung in Sinneinheiten wurde insbesondere darauf geachtet, dass diese beim Notieren für eine anschließende konsekutive Verdolmetschung jeweils ein zu notierendes bedeutungstragendes Gedankenelement in Form von Sinneinheiten bzw. Problemauslösern ausdrücken können. Für diesen Zweck wurden diese, ähnlich den Satzgliedern, in einem ersten Schritt mittels Fragepronomen (z. B. wer, was, wie viel, etc.) erkannt.

Zur besseren Übersicht wurden die Transkripte der tatsächlich vorgetragenen Reden in ein Word-Dokument mit vertikalen Zeilennummerierungen eingefügt und nach jeder erfragten Sinneinheit ein Absatz gemacht, um später mit der ebenso segmentierten Verdolmetschung verglichen zu werden.

Nachdem die Texte nach Sinneinheiten unterteilt wurden, wurden diese in eine Excel-Tabelle untereinander eingefügt, um hier die genauen Analysen der Verdolmetschungen aller Teilnehmenden jeweils in einem separaten Excel-Sheet pro Teilnehmenden, aufgeteilt auf die beiden Reden B1 und A2, vornehmen zu können.

Dabei ist zu beachten, dass die Problemauslöser, die, wie weiter unten beschrieben, kodiert und extra notiert wurden, auch die Unterteilung in Sinneinheiten beeinflussten. Es

wurde nämlich eine Unterteilung in zwei Sinneinheiten zwischen dem jeweiligen Problemauslöser und dem darauffolgenden Wort vorgenommen, falls dieses für eine weiterführende Bedeutungspräzisierung im Zuge eines neu eingeführten Fragepronomens notwendig war. Dies war insbesondere nach Zahlen, aber auch teils nach Eigennamen und Aufzählungselementen der Fall, welche eine zu präzisierende Bedeutung als Sinneinheit unter Berücksichtigung der übergreifenden Sinnübereinstimmung erforderte und kein eigenes Kompositum darstellte.

Beispiele zur Rede B1 „Gewalt gegen Frauen“:

across the 28 (1) Member States (2) sowie *the Swedish (1) cabinet (2)* im Vergleich zu *In Nordic countries (1)*

Beispiele zur Rede A2 „Erdbeben in Japan“:

Japan’s high-speed trains (1) im Vergleich zu

and how Tokyo can be affected (1) by a similar (2) or bigger earthquake (3)

Das erste Beispiel segmentierter Sinneinheiten der Rede B1 enthält eine Zahl als Problemauslöser und das zugehörige Kompositum, welches den Bezug weiter präzisiert.

Beim zweiten Beispiel wird das schwedische Kabinett als zwei Sinneinheiten gezählt, denn obwohl beide Teile auf das Fragepronomen „wer“ antworten, enthält dieses Beispiel zudem zwei Problemauslöser, einen Eigennamen und einen Terminus, welche für die detaillierte Bewertung nach Problemauslösern auch sinngemäß, aufgrund einer Präzisierung der politischen Einrichtung dank des zugehörigen Substantivs, separat betrachtet werden können.

Die nordischen Länder sind eine Sinneinheit, da etwa im Falle einer Verdolmetschung als „in Skandinavien“ bzw. „im Norden“, das heißt, wenn nur der Eigenname ohne das zugehörige Substantiv gesagt wird, nicht der allgemeine Sinn beeinträchtigt wird.

Japans Hochgeschwindigkeitszüge als erstes Beispiel zur Rede A2 werden als eine Sinneinheit gezählt, da insgesamt auf das Fragepronomen „wer/ was“ geantwortet wird und im Falle des Auslassens, also des Nicht-Verdolmetschens (0) des Problemauslösers Japan, also des Eigennamens, in der Rede zum Thema Erdbeben in Japan dank der allgemeinen Sinnübereinstimmung im Falle einer vorherigen Erwähnung weiterhin klar ist, dass über dieses Land berichtet wird.

Letztlich entstehen bei Tokio und einem ähnlichen bzw. größeren Erdbeben insgesamt drei Sinneinheiten, weil einerseits nach „wer/ was“ und einem neuen Ort bzw. in diesem Fall einer spezifischen Stadt als Eigenname gefragt wird, und andererseits entsteht die Frage nach „wovon“, aufgeteilt auf zwei leicht unterschiedliche Phänomene, nach denen gefragt werden

kann. Das möglicherweise umschriebene Verb in der ersten Sinneinheit könnte dabei keinen Bedeutungsverlust herbeiführen, wenn alle weiteren Sinneinheiten vorhanden sind. Auf diese Weise wurden die Sinneinheiten zeilenweise erkannt und sichtbar gemacht.

Spalte Fragewort

In der Spalte direkt neben der Einteilung in Sinneinheiten wurden auch die Fragepronomen eingetragen, mit deren Hilfe diese erfragt wurden. Die Sinneinheiten wurden dank des Erfragens auch insofern erkannt, als beim Auslassen einer Sinneinheit, aber beim Vorhandensein der darauffolgenden, immer noch der Gedanke des in sich abgeschlossenen Gesagten der vorherigen Sinneinheit vorhanden blieb.

Beispielsweise ergaben sich dadurch am Anfang der Rede A2 im Satz „In my previous speech (1) I talked to you about (2) (…)" zwei Sinneinheiten, denn wenn nur die Aussage darüber gedolmetscht wäre, dass in der Vergangenheit von der vortragenden Person über etwas gesprochen wurde, dann wäre der Sinn zwar verkürzt, nämlich darum, dass es eine vorherige Rede gab, aber die Information darüber, dass etwas zu einem früheren Zeitpunkt angesprochen wurde, wäre dennoch vorhanden. Das Fragepronomen wäre in diesem Fall „Wann?". Umgekehrt könnte eine vorherige Rede angesprochen werden ohne dass die vortragende Person damit in Zusammenhang gebracht wurde, wobei das zugehörige Fragepronomen in diesem Fall „Wer?" entspricht.

Spalte Problemauslöser

Rechts neben den Fragewörtern wurden dann die Problemauslöser nach Gile (1985), nämlich die Eigennamen von Personen, Orten, Institutionen, das heißt inklusive der Named Entities, die Zahlen, auch Ordinalzahlen, die Aufzählungselemente, mit mindestens drei Elementen jeweils, sowie die Termini innerhalb einer Sinneinheit erkannt und herausgeschrieben.

Wie bereits erwähnt wurde zur Beurteilung zunächst die Sinneinheit bewertet sowie dann auch weiter der Problemauslöser. Bei den Zahlen als Problemauslöser wurde bei jeglicher Ungenauigkeit eine (0) für „nicht verdolmetscht" eingetragen. Die separate Beurteilung nach Sinneinheiten führte bei einer minimalen Unterscheidung von bis zu 5 %, wie 27 statt 28 Mitgliedsstaaten bei TN3, zu einer leichten Ersetzung, bei einem Unterschied zwischen 5 bis 10 % zu einer schweren und bei Abweichungen darunter bzw. darüber zu einer kritischen Ersetzung (vgl. Spalte Fehlerkategorie).

Die englischen Eigennamen bzw. Named Entities (vgl. Abschnitt 2.2.1) sowie die Aufzählungselemente, die ineinandergriffen, etwa bei der Aufzählung von Ländern, da diese

dann sowohl als Aufzählungselemente als auch Eigennamen gezählt wurden, ließen aufgrund der punktuellen Beurteilung als Problemauslöser nur eine der unten angeführten als gültige deutsche Versionen zu, die dank des Cambridge Englisch-Deutsch-Wörterbuchs[1] bzw. dank eines Abgleichs und Erkennens weiterer passender Ausdrücke in den analysierten Verdolmetschungen ggf. um diese ergänzt wurden. Wurde der Eigenname umschrieben, etwa bei der Agentur der Europäischen Union für Grundrechte, wurde hingegen die Sinneinheit separat nach der vermittelten Bedeutung mehr oder weniger stark mittels Fehleranalyse dementsprechend nuancierter beurteilt.

Tabelle 5: Beurteilungsbasis für Eigennamen und Aufzählungselemente

Eigennamen (EN-DE)	
Rede B1: Gewalt gegen Frauen	Rede A2: Erdbeben in Japan
European Agency for Fundamental Rights – Agentur der Europäischen Union für Grundrechte	Japan – Japan (13x)
European Union – Europäische Union (EU)	Tokyo – Tokio (2x)
Nordic – nordisch, Nord-, skandinavisch (2x)	Kobe – Kobe
Southern – südlich, Süd-	Kobe Earthquake Memorial Museum – das Kobe Erdbeben-Gedenkmuseum bzw. das Kobe Earthquake Memorial Museum
Scandinavian – nordisch, skandinavisch	Japanese – japanisch
Swedish – schwedisch	Edo River – der Edo, der Fluss Edo, der Edo-Fluss
Aufzählungselemente (EN-DE)	
Denmark – Dänemark	innovate – innovativ sein
Finland – Finnland	invest - investieren
Spain – Spanien	educate – (aus)bilden, lehren
Danish – Däne, Dänin, dänisch	learn – lernen, beibringen
Rumania – Rumänien	flashlight – die Taschenlampe
	radio – der Radio

[1] abrufbar unter: https://dictionary.cambridge.org/dictionary/english-german/

	first-aid kit – der Erste-Hilfe-Kasten/ das Erste-Hilfe-Set
	enough food and water – genug/ genügend/ ausreichend Essen und Trinken

Die Verwendungshäufigkeit der Termini, welche aufgrund ihrer Themenspezifik erkannt wurden, wurde für diesen Zweck mithilfe des Oxford English Dictionary (OED) überprüft[1]. Die in den Ausgangsreden erkannten Termini wurden im Suchfeld eingegeben, und im Falle von mehrfachen Einträgen wurde jener gewählt, der dem Sinn des Ausgangstextes entsprach.

Das Wörterbuch misst die Häufigkeit anhand eines roten Balkens und fügt eine in diesem Fall irrelevante Angabe der ungefähren Jahreszahl des ersten Vorkommens dieses Terminus an. Die 8 roten Balken werden wie folgt definiert:

Tabelle 6: Häufigkeitsanzeige der Termini (in Anlehnung an: https://www.oed.com/information/understanding-entries/frequency, Datum des Zugriffs: 03.06.2024)

Balken	Häufigkeit je 1 Mio. Wörter	Beispiel
8	> 1.000-mal	the
7	100 – 1.000-mal	food
6	10 – 100-mal	tunnel
5	1 – 10-mal	flashlight
4	0,1 – 1-mal	tsunami
3	0,01 – 0,1-mal	tremor
2	< 0,01-mal	backlash
1	–	–

Als selten und terminologisch anspruchsvoll bzw. eine gezielte terminologische Vorbereitung erfordernd wurden jene Termini definiert, bei denen der Häufigkeitsbalken maximal 5 von 8 Balken lang war, da hier die Grenze zwischen einer allgemeinen Alltagssprache hin zu einer durchdachteren und letztlich stark themen- bzw. kontextabhängigen Ausdrucksweise gezogen werden konnte. Zusammengesetzte Wörter mussten separat analysiert werden und blieben daher aufgrund der Erschwernis der Zusammensetzung in der Liste enthalten, auch wenn einzelne Elemente als häufiger als die erwähnten 5 Balken bewertet wurden.

[1] abrufbar unter https://www.oed.com

Tabelle 7 zeigt die erkannten Termini mit den jeweiligen Balken, die Wörter enthält, die aufgrund ihrer geringeren Häufigkeit als Problemauslöser erachtet wurden, jeweils aufgeteilt in die beiden Reden A2 über Erdbeben in Japan und B1 über Gewalt gegen Frauen.

Da diese entsprechend „transkodiert" werden mussten (vgl. Mankauskiené 2018a: 16), wird daneben auch die im vorliegenden Fall und aufgrund des textuellen Zusammenhangs einzig gültige bzw. auf diese Version überprüfte Variante angeführt. Falls die Termini sinngemäß, aber in anderen Worten, gedolmetscht wurde, wurde zwar der Problemauslöser als nicht verdolmetscht gewertet, diese Tatsache wurde aber im Zuge der Bewertung nach Sinneinheiten in leichte, schwere und kritische Ersetzungen, Hinzufügungen und Auslassungen sehr wohl separat berücksichtigt.

Die Übersetzung wurde ebenso wie bei den Eigennamen bzw. Aufzählungselementen dank der Online-Version des Cambridge Englisch-Deutsch-Wörterbuchs[1] durchgeführt bzw. teils im Laufe der Analyse der Materialien ggf. um weitere synonyme Ausdrücke ergänzt.

Tabelle 7: Überblick über die Häufigkeit der Termini

Termini mit Häufigkeitswert (EN-DE)	
Rede B1: Gewalt gegen Frauen	
sexual harassment (3x)	die sexuelle Belästigung
egalitarian	egalitär
male-dominated.	männerdominiert/ männlich dominiert
childcare provisions	die Kinderbetreuungsangebote
cabinet (2x)	das Kabinett
patriarchal	patriarchalisch/ patriarchisch
submissive	unterwürfig/ gehorsam
to breed resentment	Ressentiments/ Unmut/ Ärger hervorrufen
backlash	der Rückschlag bzw. der Backlash
emancipation (2x)	die Emanzipation
emasculated	entmannt
urbanization (2x)	die Urbanisierung/ die Verstädterung

[1] abrufbar unter https://dictionary.cambridge.org/dictionary/english-german/

Rede A2: Erdbeben in Japan	
earthquake-resistant	erdbebenresistent/ erdbebensicher
tremor	die Erschütterung
earthquake-proof	erdbebenresistent/ erdbebensicher/ auch: Erdbebensicherheit
tsunami (2x)	der Tsunami
emergency alert system	das Notfallalarmsystem/ Notfallwarnsystem
impending	bevorstehend/ drohend
freeze	zum Stillstand bringen/ stehen bleiben
water discharge tunnel	der Wasserablasstunnel
feat	das Meisterwerk/ die Meisterleistung/ die Errungenschaft
cyclone	der Zyklon/ der Wirbelsturm

An dieser Stelle sei noch angemerkt, dass ein leicht erhöhter prozentueller Anteil an Problemauslösern in der Rede A2 zudem neben einem erhöhten kognitiven Aufwand möglicherweise auch auf einen kürzeren Abstand letzterer und somit auf eine schnellere Abfolge schließen lassen könnte, die diese um 10 wpm langsamer vorgetragene Rede (vgl. Tabelle 2) aufgrund ihrer Dichte trotzdem schneller erscheinen lassen könnte, was wiederum einen möglichen Problemauslöser ergeben kann (vgl. Seeber, 2015: 85).

Jedoch wurde dank der Berechnung des wpm bei beiden Reden erkannt, dass die Geschwindigkeit als Problemauslöser nicht nur aufgrund zahlreicher Problemauslöser der Reden an speziellen Stellen auftauchte, sondern beide Reden insgesamt beeinflusste. Obwohl die Rede A2 nämlich anteilsmäßig mehr Problemauslöser beinhaltet, ist das Tempo um 10 Wörter pro Minute geringer, was wiederum den anteilsmäßig niedrigeren Wert der Problemauslöser der Rede B1 aufgrund einer rascheren Vortragsgeschwindigkeit ausgleichen könnte.

Schließlich wurden Abkürzungen als Problemauslöser weder in der Rede B1 noch in der Rede A2 gefunden und daher entstand diese Kategorie beim Analysieren nicht.

Spalte Kodierung

Die betreffenden Segmente, die Problemauslöser enthalten, wurden in die nebenstehende Spalte eingetragen sowie mit dem Anfangsbuchstaben wie folgt eindeutig kodiert:

- E: Eigennamen inkl. Named Entities von Orten, Organisationen, etc.
- Z: Zahlen, darunter auch Ordinalzahlen
- A: Aufzählungen, wobei eine Aufzählung mindestens drei Aufzählungselemente enthielt und jedes dieser Elemente zur besseren Beurteilbarkeit als eigene Sinneinheit aufgeschlüsselt und gezählt werden sollte sowie
- T: Termini, die terminologisches Wissen bzw. eine terminologische Vorbereitung voraussetzen, da sie üblicherweise nicht oft vorkommen

Spalte Verdolmetschung

Rechts neben der Unterteilung in Sinneinheiten und dem Sichtbarmachen der darin enthaltenen Problemauslöser in der Ausgangsrede wurde der jeweilige Abschnitt in der aufgenommenen und transkribierten Verdolmetschung der jeweiligen Teilnehmenden gesucht und der Einteilung in Sinneinheiten entsprechend parallel eingefügt bzw. aligniert.

Diese Vorgehensweise erlaubte zwar aufgrund der unterschiedlichen sprachlichen Struktur zwischen dem Englischen und Deutschen, vor allem bei dem Durchgang ohne Sight-Terp, nicht immer eine eindeutige Zuordnung auf der Ebene der jeweiligen Sinneinheit (vgl. dazu die Kritik von Pöchhacker 2016: 142), hat sich jedoch als praktisch umsetzbar erwiesen, solange die Sinneinheiten in erster Linie nach ihrem unmittelbaren Vorkommen, nicht nach ihrer Position beurteilt wurden (vgl. unten *Spalte Fehlerkategorie*).

Spalte Fehlerkategorie

Die Genauigkeit wurde für die Sinneinheiten (SE) bzw. die darin enthaltenen Problemauslöser (PT) separat betrachtet, da bereits in der Unterteilung der Sinneinheiten auf eine möglichst explizite Unterscheidung in Problemauslöser und sinnverändernde (weitere) Sinneinheiten geachtet wurde.

Für die Sinneinheiten wurde die Verdolmetschung des jeweiligen Segments mit dem Original verglichen und die von Zhao (2021) (vgl. Abschnitt 2.2.3) übernommenen Fehlerkategorien in der Spalte zu den jeweiligen Abweichungen notiert und wieder kodiert, nämlich in Auslassungen (A), Hinzufügungen (H), und Ersetzungen (E). Diese wurden in einem zweiten

Schritt auf ihre Schwere hin beurteilt und mit „min“, bzw. leicht, „maj“, bzw. schwer, und „cri“, bzw. kritisch, kodiert.

Hinzufügungen und Ersetzungen drückten zunächst in ihrer leichten oder schweren Form eine zum Fragepronomen passende, wenn auch mehr oder minder stark abgeänderte, Sinneinheit aus oder änderten im Falle von Auslassungen mehr oder weniger bedeutsam, aber nicht auf kritische Art und Weise, die allgemeine Sinnübereinstimmung.

Die Kodierung SE: E(min) wäre ein Beispiel für eine leichte Ersetzung einer Sinneinheit in der Verdolmetschung, beispielsweise wenn statt „in den Mitgliedsstaaten der EU“ „in Ländern der EU“ gesagt wurde (vgl. TN1), denn dann wurde auf die Frage „wo“, also auf die Sinneinheit, die sich auf einen Ort bezieht, geantwortet, jedoch wurde ein (passender) Überbegriff eingesetzt. Wenn aber im Gegensatz dazu als Antwort auf die Frage „wer“ und „was“ jeweils statt „Frauen wurden sexuell belästigt“ „Frauen wurden vergewaltigt“ (vgl. TN1) gedolmetscht wurde, so wäre diese Ersetzung in der Antwort auf „was“ als SE: E(maj) schwer, nicht zuletzt deshalb, da der Sinn verzerrt wurde.

Zudem konnten Auslassungen unbedeutsam sein, wenn sie davor bereits ggf. in einer anderen Form genannt wurden. Es war aus diesem Grund auch nicht immer eindeutig eine Grenze zwischen schweren und leichten Fehlerkategorien zu ziehen. Beispielsweise verdolmetschte TN1 statt „sexueller Belästigung“ „Gewalt“; hier galt es dann neben der vorhandenen Antwort auf das relevante Fragepronomen abzuwägen, ob der Sinn bei dieser Ersetzung grundsätzlich gegeben ist. Dieser Fall wurde aufgrund eines Zusammenhängens beider Phänomene mit SE: E(min), also einer leichten Ersetzung, markiert. Ebenso musste bei der Auslassung von Antworten auf Fragepronomen entschieden werden, ob das (teilweise) (Nicht-) Vorhandensein die Aussage als Gesamtes leicht oder stark verändert. Falls die Information zwar nicht dem übergreifenden Verständnis dienlich war, aber in der Verdolmetschung komplett fehlte, wurde diese als schwerer Fehler markiert. Falls die Information in der Verdolmetschung in einer anderen, zum Ausgangstext passenden Form davor oder danach genannt wurde, war die Auslassung an der jeweiligen Stelle, an der sie im Ausgangstext vorkam und an der sie für die Sinnübereinstimmung relevant war, lediglich leicht.

Genauso wurde auch darauf geachtet, dass nur solche als kritische Hinzufügungen galten, die in der Rede nicht genannte und sinnverändernde Informationen enthielten, nicht etwa Hinzufügungen, die möglicherweise bewusst aufgrund der mündlichen Natur der Verdolmetschung eingesetzt wurden. Hinzufügungen, die tatsächlich Gesagtes sinngemäß ergänzten, bzw. dementsprechende Wiederholungen, wurden als leicht, Hinzufügungen, die neue, aber

den übergeordneten Sinn nicht im Grunde verzerrende Informationen einführten, wurden als schwer gewertet.

Letztlich nahmen kritische Fehlerkategorien (cri) gar nicht auf das Fragepronomen bzw. die jeweilige Sinneinheit Bezug bzw. beeinträchtigten die allgemeine Sinnübereinstimmung. Es wurden offensichtlich sinnverändernde Elemente hinzugefügt oder stark abgeändert bzw. ersetzt oder hielten im Falle von Auslassungen für das gegebene Thema sachdienliche Informationen komplett zurück, die auch in einer anderen Form nirgendwo genannt wurden. Im obigen Beispiel zu den Mitgliedsstaaten der EU wäre demnach ein komplettes Auslassen oder das Ersetzen durch ein Segment, das sich auf ein anderes Fragepronomen als auf „wo" bezieht, eine kritische Auslassung. Ein weiteres Beispiel für eine kritische Ersetzung von TN2 wäre die Verdolmetschung der Aussage „Ein weiterer Faktor, der damit einhergeht, ist die Urbanisierung", wobei diese in der Ausgangsrede gegenteilig als „unconnected factor" erwähnt wurde und somit den Sinn kritisch verändert.

Der Beurteilungsfokus lag dabei stets auf einer Sinnübereinstimmung mit der ausgangssprachlichen Äußerung (vgl. Bühler, 1986: 233) sowie auf dem Erkennen sinn- bzw. aussageverändernder Segmente. Das bedeutet, dass in Anlehnung an die Bewertungsmethode von Gieshoff & Albl-Mikasa (2022: 213) Abweichungen wie Verallgemeinerungen, grammatikalische Fehler und eine geänderte Reihenfolge der Sinneinheiten bzw. Problemauslöser nicht einer kritischen Fehlerkategorie zugeordnet wurden, solange sie für die Aussage bzw. die Bedeutung der Sinneinheit nicht entscheidend waren und diese nicht grob verzerrten.

Insbesondere da eine vertikal nebeneinander angeordnete Entsprechung der Sinneinheiten bzw. der von den Problemauslösern beeinflussten ausgangssprachlichen und zielsprachlichen Segmentierung aufgrund der unterschiedlichen sprachlichen Struktur zwischen dem Englischen und dem Deutschen nicht immer möglich war, war zusätzlich eine geänderte Reihenfolge, falls diese nicht für die logische Abfolge entscheidend war, keine zu zählende Fehlerkategorie. Etwa wurde erkannt, dass bei den Durchgängen ohne Sight-Terp nicht strikt die ursprüngliche Reihenfolge eingehalten wurde, was aber nicht zwangsläufig zu Unverständnis führt, sondern auch eine mündlich freiere, das heißt weniger am Ausgangstext orientierte Ausdrucksweise auszeichnen kann, die hier nicht der Beurteilungsgegenstand war. Falls eine Sinneinheit zu einem früheren bzw. ggf. späteren Zeitpunkt als an der Stelle im Ausgangstext genannt wurde und die Reihenfolge nicht für die Sinnübereinstimmung entscheidend war, so wurde in der entsprechenden Zeile unter Angabe der Zeilennummer darauf verwiesen und daher musste das jeweilige Segment nicht als Auslassung bzw. Hinzufügung markiert werden.

Für Problemauslöser bzw. Named Entities wie Zahlen, Eigennamen von Orten bzw. Institutionen sowie Aufzählungen und die verwendeten Termini wurde hingegen nur eine, nämlich zumeist die offizielle Version als vorhanden bewertet. Beispielsweise galten Umschreibungen der Agentur der Europäischen Union für Grundrechte (etwa europäische Agentur für grundlegende Menschenrechte bei TN1) ebenso wie eine anderweitige Veränderung von offiziellen Institutions- bzw. Ortsbezeichnungen (etwa der Fluss Itu statt Edo bei TN1) als nicht verdolmetschter Problemauslöser, aber in Hinblick auf die Beurteilung der Sinneinheit lediglich als eine Form von Ersetzung. Die binäre Einteilung in die Kategorien „verdolmetscht" (1) bzw. „nicht verdolmetscht" (0) der analysierten Einheiten wurde zum Zwecke einer punktuellen Bewertung gewählt (vgl. Abschnitt 2.2.1).

Für beide Kategorien galt, dass wenn zuerst der korrekte und direkt danach ein falscher Problemauslöser bzw. Sinneinheit gesagt wurde, die letzte als gültige Version und somit als korrekte oder fehlerhafte Verdolmetschung gezählt wurde.

Dass die prosodischen Merkmale hier nicht analysiert wurden, sei nochmals kritisch angemerkt, da eine Bewertung einer Verdolmetschung in ihrer Gesamtheit, wie im Theorieteil unter Kapitel 2 angesprochen, sehr wohl mehr als nur das Zählen der Fehler umfasst. Aus diesem Grund wurde zumindest wie oben erwähnt darauf geachtet, dass die Sinneinheiten, welche in sich abgeschlossen eine Bedeutung ausdrücken, nicht auf Wortebene analysiert wurden. Beispielsweise verdolmetschte TN5 folgendermaßen aus dem Englischen ins Deutsche:

Original: So, when I tried to think about the reasons behind this strange result (…)
TN5: Was könnten die Gründe sein, dass vor allem in den nordischen Ländern diese Zahlen so hoch sind?

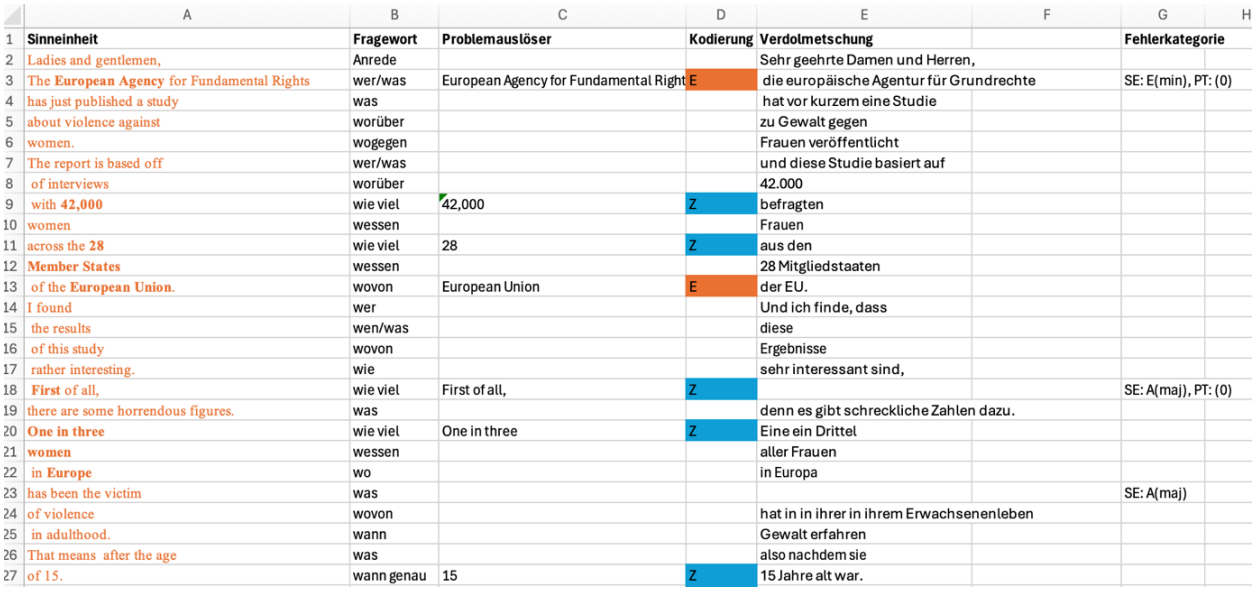
Die Verdolmetschung der Sinneinheiten entspricht zwar nicht wörtlich dem Original, aber die Bedeutung, dass über die Gründe hinter der zuvor angesprochenen paradoxen Thematik des höheren Gewaltaufkommens in den nordischen Ländern nachgedacht wird, ist sehr wohl korrekt vermittelt worden und daher wurde dieser Fall nicht als Fehler markiert.

Die Analyse der Sinneinheiten erlaubte somit eine breitere Sichtweise, da sie neben den Problemauslösern, die als „verdolmetscht" (1) oder „nicht verdolmetscht" (0) gewertet wurden, weil z. B. Eigennamen zumeist nur eine korrekte Lösung zulassen, auch eine sinnbasierte Metaebene zuließen.

Das bedeutet, dass „one in three“ und „women“ als zwei Sinneinheiten gelten, da Zahlen als Problemauslöser neben der punktuellen Analyse nach den Kategorien „verdolmetscht“ (1) oder „nicht verdolmetscht“ (0) aufgrund des Fragepronomens „wie viele“ als Sinneinheit auch auf der Sinnebene analysiert werden können. Wenn nämlich im vorliegenden Beispiel statt des genauen Anteils „ein Drittel“ lediglich „ein Teil“ gedolmetscht wird, dann wäre einerseits der Sinn schwer beeinträchtigt, andererseits wäre der Problemauslöser nicht verdolmetscht. Besonders nach Zahlen als Problemauslöser macht es außerdem einen kritischen Bedeutungsunterschied, wenn im darauf Bezug nehmenden Substantiv, erfragt mit „wessen“, hier statt Frauen beispielsweise Männer oder gar nichts gesagt wird, weshalb hier eine Trennung vorgenommen werden musste.

Zu den Aufzählungen, die insbesondere in der Rede A2 über Erdbeben in Japan vorkamen, sei abschließend angemerkt, dass diese zwar im Falle einer Auslassung als nicht verdolmetschte Problemauslöser gezählt wurden, die Sinneinheit wurde beim Abändern, wenn etwa „als Nächstes“ anstelle der Ordinalzahl und des zugehörigen Substantivs gesagt wurde, nur als leichte Auslassung, beim kompletten Weglassen als schwere, nicht aber als kritische Auslassung gewertet, da der übergeordnete Sinn bzw. die Struktur aufgrund weiterer, u. a. prosodischer Merkmale, ersichtlich war. Der Einfluss der prosodischen Merkmale auf die Verdolmetschung wurde zwar nicht separat untersucht, aber eine dementsprechend weniger strenge Beurteilung berücksichtigte zum Teil deren Einsatz.

Schlussendlich wurden auf diese Weise aus den je zwei Durchgängen pro Teilnehmenden insgesamt 20 Audioaufnahmen der Verdolmetschungen transkribiert, analysiert und wie oben beschrieben bewertet. Die ausgefüllten Bewertungsbögen wurden im Anschluss innerhalb der beiden Durchgänge als auch übergreifend mit den Ergebnissen anderer Teilnehmenden und Reden quer verglichen. Darauf aufbauend konnten Grafiken erstellt und Aussagen über die Verdolmetschungen mit bzw. ohne Softwareanwendung sowie in Hinblick auf den Schwierigkeitsgrad der Reden getroffen werden.



	A	B	C	D	E	F	G	H
1	**Sinneinheit**	**Fragewort**	**Problemauslöser**	**Kodierung**	**Verdolmetschung**		**Fehlerkategorie**	
2	Ladies and gentlemen,	Anrede			Sehr geehrte Damen und Herren,			
3	The **European Agency** for Fundamental Rights	wer/was	European Agency for Fundamental Right	E	die europäische Agentur für Grundrechte		SE: E(min), PT: (0)	
4	has just published a study	was			hat vor kurzem eine Studie			
5	about violence against	worüber			zu Gewalt gegen			
6	women.	wogegen			Frauen veröffentlicht			
7	The report is based off	wer/was			und diese Studie basiert auf			
8	of interviews	worüber			42.000			
9	with **42,000**	wie viel	42,000	Z	befragten			
10	women	wessen			Frauen			
11	across the **28**	wie viel	28	Z	aus den			
12	**Member States**	wessen			28 Mitgliedstaaten			
13	of the **European Union**.	wovon	European Union	E	der EU.			
14	I found	wer			Und ich finde, dass			
15	the results	wen/was			diese			
16	of this study	wovon			Ergebnisse			
17	rather interesting.	wie			sehr interessant sind,			
18	**First** of all,	wie viel	First of all,	Z			SE: A(maj), PT: (0)	
19	there are some horrendous figures.	was			denn es gibt schreckliche Zahlen dazu.			
20	**One in three**	wie viel	One in three	Z	Eine ein Drittel			
21	**women**	wessen			aller Frauen			
22	in **Europe**	wo			in Europa			
23	has been the victim	was					SE: A(maj)	
24	of violence	wovon			hat in in ihrer in ihrem Erwachsenenleben			
25	in adulthood.	wann			Gewalt erfahren			
26	That means after the age	was			also nachdem sie			
27	of 15.	wann genau	15	Z	15 Jahre alt war.			

Abbildung 7: Bildschirmaufnahme des Analyse- und Bewertungsrasters in Excel von TN8

Abbildung 7 zeigt zur Veranschaulichung der Vorgehensweise einen Teil des ausgefüllten Analyserasters von TN8 zur Rede über Gewalt gegen Frauen, der in die oben erläuterten sechs Spalten eingeteilt und auf diese Weise eine Bewertung der beiden Verdolmetschungen untereinander zuließ.

Beim Befüllen des Analyserasters fiel auf, dass der Durchgang, bei dem die Verdolmetschung mit Sight-Terp stattfand, leichter zu alignieren war als jener ohne Sight-Terp. Die Struktur wurde nämlich oft gut erkennbar beibehalten und entsprach dem Ausgangstext. Trotzdem wurde aufgrund der allgemeinen Sinnübereinstimmung eine geänderte, aber verständliche Chronologie, wie bereits beschrieben, nicht als Fehler gewertet.

Im Folgenden wird die Vorgehensweise des Leitfaden-Interviews erläutert, bevor dann eine Überleitung zur Präsentation der Ergebnisse des Experiments und des Interviews erfolgt.

4.3 Leitfaden-Interview

Leitfaden-Interviews sind Interviews, die den Ablauf eines Gesprächs mithilfe eines vorbereiteten Leitfadens gestalten (Helfferich, 2014: 560). Die Vorgehensweise erfolgt systematisch und kann sowohl Aufforderungen, vorformulierte Fragen, Stichworte sowie Vereinbarungen für die Handhabung spezieller Phasen des Interviews beinhalten, wobei eine maximale Offenheit aus pragmatischen Gründen eingeschränkt werden kann, um ein gewisses Maß an Struktur bzw. Steuerung zu erhalten (Helfferich, 2014: 560).

Als Mittel zur qualitativen Datengewinnung wurde im Rahmen dieses empirischen Teils das Leitfaden-Interview gewählt, da das computergestützte Konsekutivdolmetschen

aufgrund seiner Neuheit (vgl. Kapitel 3) einen möglichst offenen Raum für Diskussions- und Reflexionsargumente erforderte, an die in vorab formulierten Fragen eventuell nicht gedacht werden konnte. Dadurch konnten spannende Einblicke in die Arbeitsweise und in die Gedanken der Teilnehmenden gewährt werden, die auch für den künftigen Technologieeinsatz beim konsekutiven Dolmetschen relevant sein könnten.

Es wurde dabei zudem darauf geachtet, dass der Redefluss während des Gesprächs möglichst erhalten bleibt und ggf. wurden beim Aufkommen spannender Gedanken weitere Folgefragen gestellt, die vom „Skelett" des ursprünglichen Leitfaden-Interviews abwichen. Jedoch zogen sich die vorab ausformulierten Fragen wie ein roter Faden durch das Gespräch, insbesondere falls dieses abzureißen drohte.

4.3.1 Vorgehensweise

Nach der Durchführung der beiden Verdolmetschungen wurde die Audioaufnahme am Diktiergerät nicht beendet, sondern lief weiter, denn den Teilnehmenden wurden im Rahmen eines Leitfaden-Interviews Fragen in Bezug auf ihre Arbeit mit Sight-Terp sowie zur Selbsteinschätzung der Verdolmetschung jeweils mit und ohne dem Tool und generell zu ihren Gedanken und Überlegungen zur offenen Diskussion gestellt.

Dabei wurde in dem hier durchgeführten Leitfaden-Interview darauf Wert gelegt, dass die Fragen zuerst grob und dann genauer in Subfragen unterteilt wurden, um den Teilnehmenden zunächst einen möglichst breiten Spielraum für die Ausformulierung eigener Gedanken zu erlauben, um sie in ihren Ideen also nicht einzuschränken, und um andererseits im Falle von starken thematischen Abschweifungen oder Gedankenlücken verständlich zu machen, worauf die Frage genau abzielt bzw. welche Erinnerungen aktiviert werden sollten.

Die Ausformulierung der Fragen war dabei in erster Linie von den Forschungsfragen dieser Arbeit inspiriert, da die Teilnehmenden vor allem ihre Arbeitsweise und die komparative Bewertung der Verdolmetschung mit bzw. ohne Sight-Terp kommentieren sollten (vgl. unten Abschnitt 4.3.2). Mit den Interviews wurde versucht auf die Thematik des computergestützten Konsekutivdolmetschens eine breitere, da plurale, Sichtweise zu erschließen sowie mögliche Verbesserungsvorschläge und Zukunftspotenziale das Tool betreffend einzuholen.

4.3.2 Aufbau des Leitfaden-Interviews

Der Leitfaden für das nach dem Experiment durchgeführte Gespräch umfasste als Grundelemente nachstehende Fragestellungen, wobei unter jeder Hauptfrage auch Subfragen

aufgelistet wurden, die die Hauptfrage genauer definierten, um den Teilnehmenden die Fragen so präzise wie möglich zu stellen.

Außerdem sei angemerkt, dass die offene Gestaltung der Diskussion dazu führte, dass bei den jeweiligen Teilnehmenden, je nach Entwicklung des Gesprächsverlaufs, teilweise auch weitere, individuelle Fragen gestellt wurden, an die bei der Vorbereitung der Fragen noch nicht gedacht werden konnte und die das Leitfaden-Interview somit im Gespräch weiterentwickelten. Diese befinden sich zwar nicht in den hier vordefinierten Fragen, erlauben aber sehr wohl im Analyseteil zum Leitfaden-Interview wertvolle Einblicke zu weiteren relevanten Aspekten.

Letztendlich wurde die in Tabelle 8 präsentierte Struktur jedoch als roter Faden verwendet, der sich sichtbar durch das Gespräch zog und den Gesprächsverlauf lenkte. Die Fragen sind in der nachstehenden Tabelle aufgelistet und rechts kurz erläutert, um verständlich zu machen, welche Informationen abgefragt wurden bzw. worauf die Fragestellung abzielte.

Tabelle 8: Erläuterung der Fragen des Leitfaden-Interviews

Personenbezogene Daten: Name, Matrikelnummer, TN-Identifikation (z. B. TN1, etc.), Semester, Studienschwerpunkt, Sprachenkombination	Diese Informationen waren für die Profilerstellung der Teilnehmenden von Bedeutung.
Hauptfrage 1: Wie schwierig empfanden Sie die Ausgangsreden? **Subfragen:** Wie schwierig waren die Reden sowohl inhaltlich als auch von der Art der Präsentation her (Geschwindigkeit, Intonation, etc.)? Hatten Sie Vorwissen (bspw. dank der Vorbereitung, Glossarerstellung) zu den behandelten Themen, auf das Sie zurückgreifen konnten?	Diese Frage wurde gestellt, um die Teilnehmenden nach ihrer Einschätzung des empfundenen Schwierigkeitsgrades der jeweiligen Reden zu fragen sowie, um mögliche Einflussfaktoren auf die Verdolmetschung, wie die Vorbereitung und das mitgebrachte Vorwissen, aufzudecken.
Hauptfrage 2: Wie empfanden Sie das Arbeiten mit Sight-Terp?	Hier wurde grundlegend nach der Art gefragt, wie die Teilnehmenden mit Sight-Terp arbeiteten. Es war wichtig herauszufinden, wie sie

Subfragen: Ist es einfach zu verstehen und zu verwenden? Ist es praktisch und nützlich oder eher ablenkend und unverlässlich? Würden Sie es bei einem realen Einsatz verwenden wollen?	bewusst mit dem Tool während der konsekutiven Verdolmetschung umgingen, da sich dieser Aspekt auch in der Forschungsfrage der vorliegenden Arbeit wiederfindet. Außerdem war es als weiterer Punkt interessant zu hinterfragen, ob sie sich eine Verwendung von Sight-Terp im Rahmen eines realen Dolmetschauftrags vorstellen könnten.
Hauptfrage 3: Wofür haben Sie Sight-Terp in erster Linie verwendet? **Subfragen**: Wofür ist Sight-Terp nützlich(er), auch im Vergleich zur klassischen Notizentechnik? (z. B. für ein erstes Verständnis, für die Schaffung eines Überblicks über Eigennamen, Zahlen, etc., zur Kontrolle der eigenen Notizen, falls diese parallel dazu gemacht wurden, für das Vom-Blatt-Übersetzen der durch Sight-Terp transkribierten Ausgangsrede bzw. für die (post-editierte) Wiedergabe der maschinellen Übersetzung in der Zielsprache, etc. …)	Diese Frage zielte auf eine wesentlich genauere Beschreibung der Anwendungstechnik von Sight-Terp ab, da die vorherige Frage, entsprechend ihrer Formulierung, zumeist nur eine erste, überblicksmäßige Antwort auf die Arbeit mit dem Tool lieferte. Die Teilnehmenden wurden gebeten, gedanklich zu rekonstruieren, wie sie in den jeweiligen kognitiven Phasen (vgl. Abschnitt 1.2.1), etwa des Zuhörens, Mitnotierens, Mitlesens und der Verdolmetschung, jeweils mit Sight-Terp bzw. ihren Notizen umgegangen sind. Die in Klammer aufgelisteten Anwendungsmöglichkeiten wurden auszugsweise genannt, um den Teilnehmenden Ansatzpunkte für das Zurückerinnern an die Arbeitsweise mit Sight-Terp zu liefern.
Hauptfrage 4: Wie schätzen Sie die eigene Leistung bzw. Qualität mit Sight-Terp im Vergleich zur Leistung mit der klassischen Notizentechnik selbst ein?	Diese introspektive Frage forderte die Teilnehmenden auf, ihre Verdolmetschungen ad hoc selbst zu vergleichen und nach ihrer Leistung zu beurteilen, wobei nicht genau gesagt wurde, wie sie Leistung bzw. Qualität (vgl. Kapitel 2) definieren sollten.

	Manche TN definierten selbst die Leistung als solches genauer, etwa bezüglich der Genauigkeit und Flüssigkeit, und kommentierten diese dementsprechend.
Hauptfrage 5: Gibt es etwas, das Sight-Terp besser oder anders machen sollte?	Diese abschließende Frage diente dazu herauszufinden, ob es bei der Arbeit mit Sight-Terp Störfaktoren gab und welche diese waren bzw. auch, um Verbesserungsvorschläge für die Weiterentwicklung des Tools bzw. generell für diesen Arbeitsmodus zu liefern.
Hauptfrage 6: Gibt es noch etwas, das Sie hinzufügen möchten?	Hier wurde den Teilnehmenden die Chance gegeben, mögliche Gedanken auszudrücken, die nicht explizit im Zuge der gestellten Hauptfragen zur Sprache kamen.

4.3.3 Analysemethode

Die Analyse der qualitativen Materialien, die im Zuge des Leitfaden-Interviews gesammelt wurden, erfolgte mithilfe der von Mayring (2022: 50) beschriebenen Technik im Rahmen der qualitativen Inhaltsanalyse. Dabei werden die im Rahmen eines Gesprächs erhobenen Informationen in zuvor gebildete Kategorien eingeteilt und so strukturiert.

Die Kategorienanwendung nach Mayring (2022: 96) erfolgt demnach deduktiv, sodass das Kategoriensystem zunächst klar definiert sowie aus der Fragestellung abgeleitet und theoretisch begründet wird (vgl. Abbildung 8). Für die Zuordnung eines Segments in eine Kategorie muss zunächst die Kategorie definiert werden. Dann werden Beispiele genannt, die sogenannte Prototypen für diese Kategorie darstellen, und schließlich werden klare Regeln da formuliert, wo es zu Problemen der Abgrenzung zwischen zwei Kategorien kommen kann, um Abgrenzungsschwierigkeiten möglichst zu vermeiden (Mayring, 2022: 96).

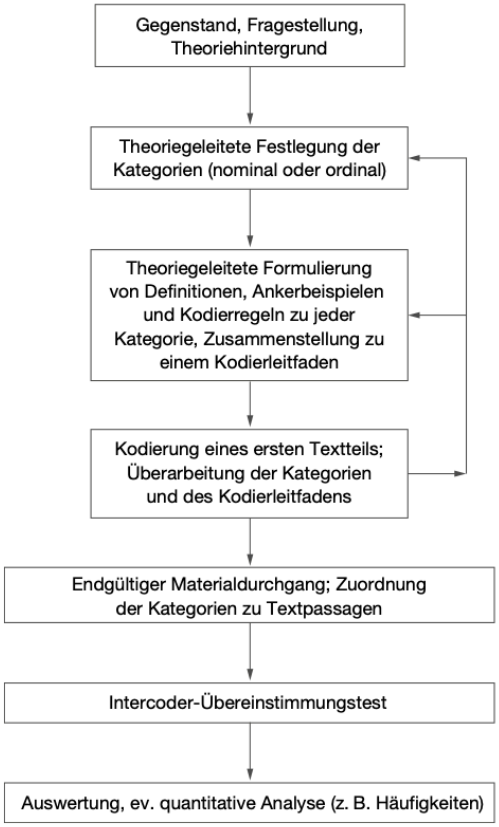


Abbildung 8: Erklärung der Kategorienanwendung (Mayring, 2022: 97)

Kategorie	Definition	Ankerbeispiele	Kodierregeln
K1: hohes Selbstvertrauen	Hohe subjektive Gewissheit, mit der Anforderung gut fertig geworden zu sein, d. h. ● Klarheit über die Art der Anforderung und deren Bewältigung ● positives, hoffnungsvolles Gefühl beim Umgang mit der Anforderung ● Überzeugung, die Bewältigung der Anforderung selbst in der Hand gehabt zu haben.	»Sicher hat's mal ein Problemchen gegeben, aber das wurde dann halt ausgeräumt, entweder von mir die Einsicht, oder vom Schüler, je nachdem, wer den Fehler gemacht hat. Fehler macht ja ein jeder.« (17, 23) »Ja klar, Probleme gab's natürlich, aber zum Schluss hatten wir ein sehr gutes Verhältnis, hatten wir uns zusammengerauft.« (27, 33)	Alle drei Aspekte der Definition müssen in Richtung »hoch« weisen, es soll kein Aspekt auf nur mittleres Selbstvertrauen schließen lassen Sonst Kodierung »mittleres S.«

Abbildung 9: Beispiel eines Kodierleitfadens (Mayring, 2022: 100)

Es entsteht somit ein Kodierleitfaden. Abbildung 9 macht ersichtlich, dass es beim Kodierleitfaden zunächst der Schaffung einer Kategorie bedarf. Daran anschließend folgen deren Definition, die theoriegestützt bzw. aus der wissenschaftlichen Literatur erklärbar ist, sowie Ankerbeispiele, die aussagekräftige Beispiele für diese Kategorie darstellen und als Referenz angesehen werden können, sowie eindeutige Kodierregeln, um Abgrenzungsprobleme und das Auftauchen von Mischkategorien zu minimieren. Diese Vorgehensweise erwies sich als sehr hilfreich für die vorliegende Arbeit.

Der im Rahmen der qualitativen Inhaltsanalyse angewandte Kodierleitfaden ist im Anhang III dargestellt. Zusammenfassend kann gesagt werden, dass die in Tabelle 8 erläuterten Hauptfragen im Kodierleitfaden die Kategorien K1 bis K6 bildeten und entsprechende Aussagen mit einer Farbe kodiert wurden. Die Definition, Ankerbeispiele und Kodierregeln dienten einer strukturierten Vorgehensweise.

Bevor der Kodierleitfaden angewandt werden konnte, mussten alle Aufnahmen transkribiert werden. Dies geschah mit der Diktierfunktion von Microsoft Word. Das somit entstandene Transkript musste korrigiert werden. Dies geschah nach den Transkriptionsregeln von Kuckartz & Rädiker (2022: 200), welche zusammenfassend Folgendes besagen und auch dahingehend angewandt wurden:

1) Jeder Sprachbeitrag bildet einen eigenen Absatz
2) Die Moderator*in und die befragten Personen werden mit einem Kürzel identifiziert
3) Es soll wörtlich transkribiert werden
4) Es wird standardsprachlich notiert
5) Deutliche Pausen werden markiert
6) Betonungen werden hervorgehoben
7) Lautes Sprechen wird durch Großbuchstaben gekennzeichnet
8) Lautäußerungen werden nicht mit transkribiert
9) Nur inhaltlich relevante Fülllaute werden transkribiert
10) Störungen von außen werden transkribiert
11) Lautäußerungen, wie Lachen, werden in einfachen Klammern notiert
12) Nonverbale Kommunikation wird auch in einfachen Klammern notiert
13) Unverständliche Wörter werden durch (unv.) sichtbar gemacht
14) Zeitmarken werden am Ende jedes Beitrags eingefügt

Insbesondere die Punkte eins bis vier, neun und vierzehn waren für die vorliegende Transkription relevant bzw. traten häufig auf. Die Moderatorin wurde mit M abgekürzt, die interviewte Person jeweils mit der zugehörigen Teilnehmenden-Identifikation TN plus Zahl.

Da die Daten computergestützt ausgewertet wurden, wurde die kostenlose Testversion des eigens für die qualitative Inhaltsanalyse von Udo Kuckartz entwickelten Programms MAXQDA[1] verwendet (Mayring, 2022: 110f.). Die Testversion wurde am 14.06.2024 für einen Zeitraum von 14 Tagen aktiviert. Nachdem die kostenlose Testversion heruntergeladen wurde, konnte mithilfe eines hilfreichen Online-Tutorials zu den ersten Schritten mit MAXQDA[2] die Arbeitsweise selbst erlernt und sofort angewandt werden.

[1] abrufbar unter https://www.maxqda.com/de/
[2] abrufbar unter https://www.maxqda.com/de/maxqda-video-tutorials

Dafür wurde jeweils ein Word-Dokument für die bereits vorab transkribierten und oben beschriebenen Audioaufnahmen der Teilnehmenden erstellt und in das System geladen. Der gesamte Text konnte analysiert werden, indem mithilfe von Farben bzw. Codes eindeutige Kodierungen zu jeder der oben präsentierten Fragen erstellt wurden. Jeder Code erhielt zudem eine Erklärung darüber, welche Inhalte in diese Kategorie fallen. Die Codes entsprechen den Kategorien laut dem oben präsentierten Kodierleitfaden und orientierten sich sehr stark an den gestellten Fragen und den erhofften Einblicken (vgl. Abschnitt 4.3.2).

Danach konnte mit der Analyse begonnen werden, indem die Transkripte gelesen und an den entsprechenden Stellen die vordefinierten Codes in den Redebeiträgen der Teilnehmenden wiedererkannt und angewandt wurden. Es entstand ein strukturierter Text, mit einem Klick auf die entsprechende Markierung wurde das Textsegment in der Code-Farbe markiert und klar kenntlich gemacht.

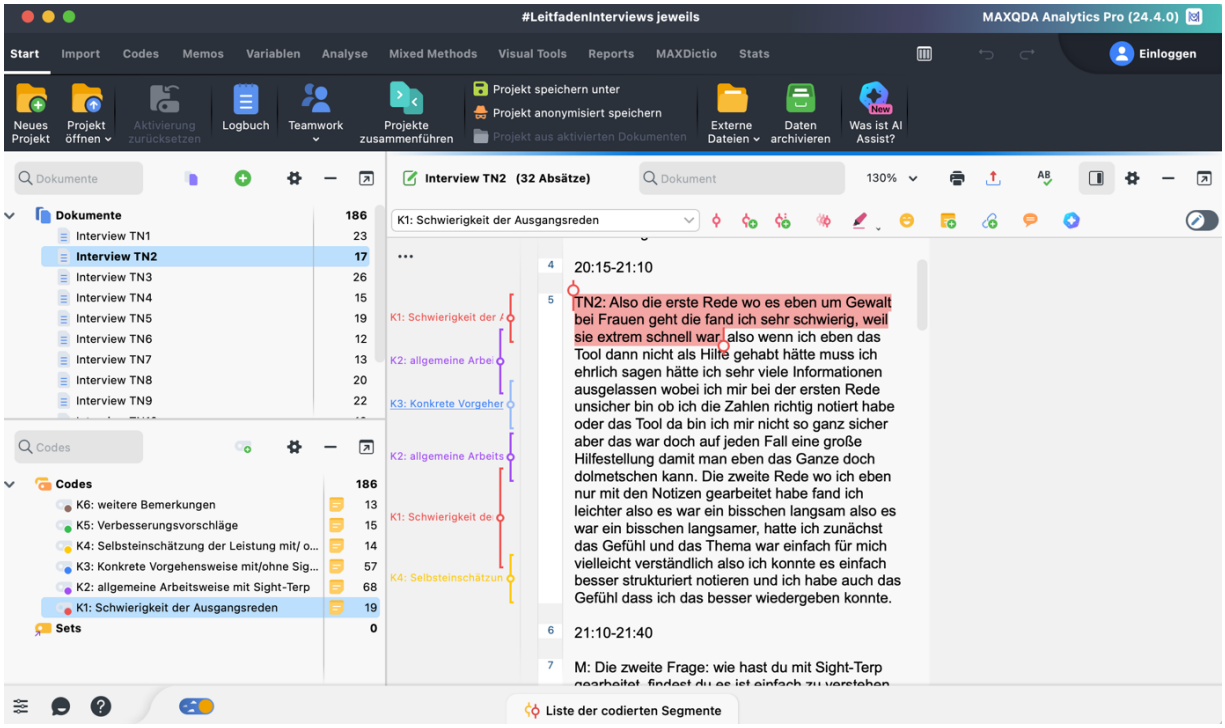


Abbildung 10: Benutzeroberfläche von MAXQDA mit einem Beispiel von TN2

Abbildung 10 zeigt die Benutzeroberfläche der heruntergeladenen Testversion des Programms MAXQDA anhand eines Ausschnitts aus dem Leitfaden-Interview mit TN2. In der linken Spalte unter dem Reiter „Dokumente“ sind alle Transkriptionen der Leitfaden-Interviews mit den entsprechenden Teilnehmenden hinterlegt und mit „Interview TN1“, „Interview TN2“, etc. benannt. Mit einem Klick auf das jeweilige Teilnehmenden-Interview kann das Dokument geöffnet und analysiert werden, indem im Text Antworten auf die kategorisierten und mit Code versehenen Fragen erkannt und dem jeweiligen Code per Mausklick zugeordnet werden. Die den Farbcodes zugeteilten Kategorien K1 bis K6 befinden sich im linken unteren Bereich. Das Zuordnen zu Kategorien ermöglicht am Ende eine Aufstellung und einen Export der nach jeweiliger Fragenkategorie sortierten Antworten aller Teilnehmenden durch einen Klick auf den gewünschten Code. Diese Vorgehensweise führt somit zu einer zeiteffizienten und produktiven Analyse der Antworten auf individueller Basis je Teilnehmenden, aber auch als Gesamtüberblick je Fragestellung.

Die Ergebnisse der Analysen sowohl der Verdolmetschungen, die im Zuge des Experiments durchgeführt wurden, als auch der aufgezeichneten Antworten im Rahmen des Leitfaden-Interviews werden im folgenden Kapitel für alle Teilnehmenden separat präsentiert.

5. Ergebnisse

5.1 Experiment

TN1

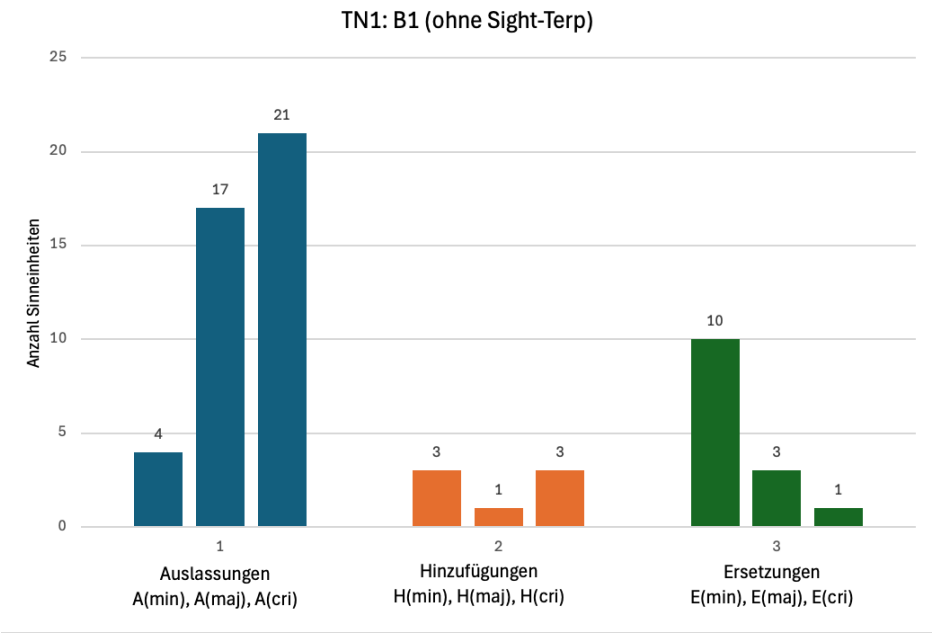


Abbildung 11: Analyseergebnisse von TN1 (ohne Sight-Terp)

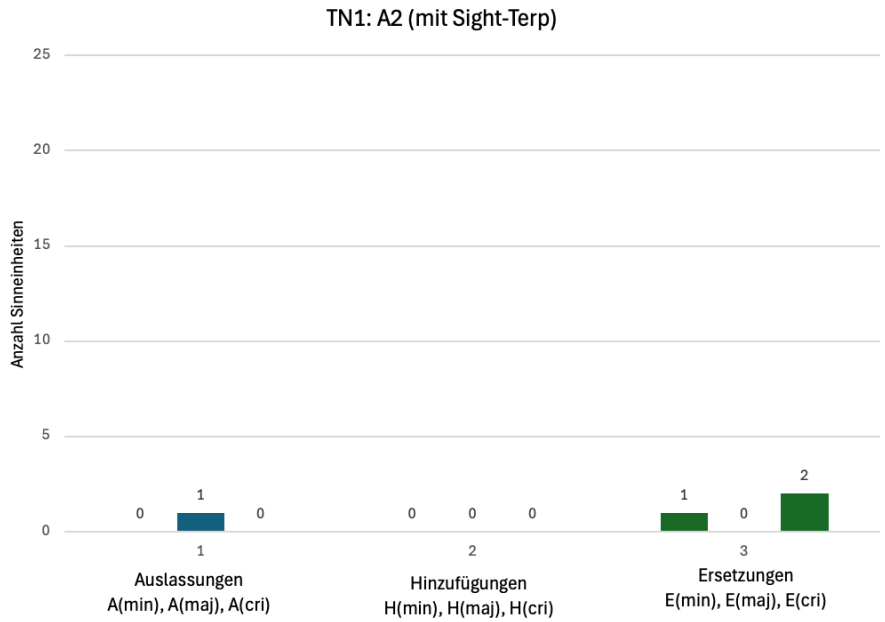


Abbildung 12: Analyseergebnisse von TN1 (mit Sight-Terp)

Die Analyseergebnisse von TN1 sind in den Abbildungen 11 und 12 ersichtlich. Dabei ist zu erkennen, dass bei der Rede B1, die ohne Sight-Terp gedolmetscht wurde, in allen drei Fehlerkategorien, also leichte, schwere und kritische Auslassungen, Hinzufügungen und Ersetzungen, mehr Fehler auftraten als bei der Verdolmetschung der Rede A2 mit Sight-Terp.

Die Fehlerquote wurde ebenso berechnet, indem die Anzahl der fehlerhaft gedolmetschten Sinneinheiten bzw. Problemauslöser jeweils durch die Gesamtanzahl der Sinneinheiten bzw. Problemauslöser in der gegebenen Rede dividiert wurde. Diese beträgt für die Rede B1 39,6 % für die Sinneinheiten und 37,5 % für die Problemauslöser sowie für die Rede A2 in Bezug auf die Sinneinheiten knapp 3 % sowie bezogen auf die Problemauslöser 0 %. Das bedeutet, dass die Verdolmetschung der Sinneinheiten um 36,7 % und jene der Problemauslöser um 37,5 % dank der technologischen Unterstützung von Sight-Terp verbessert werden konnte.

Besonders hervorzuheben ist bei TN1 außerdem, dass die kritischen Fehlerkategorien, also jene, die die allgemeine Sinnübereinstimmung stark beeinträchtigten bzw. änderten, bis auf zwei Ersetzungen, nämlich „jeden fahrenden Zug im Land einzufrieren“ für „to freeze every moving train“ sowie „Itu Fluss“ für „Edo River“, dank Sight-Terp so gut wie ausgeschlossen werden konnten. Die größte Herausforderung in der Rede B2, die ohne Sight-Terp gedolmetscht wurde, betraf die Vollständigkeit, da hier viele schwere und leichte Auslassungen in der Analyse erkannt wurden. Schließlich ist zu erwähnen, dass alle Problemauslöser in der Rede A2 mit Sight-Terp korrekt wiedergegeben wurden.

TN1 fertigte die Notizen für die Rede B1 ohne CAI-Tool auf zwei Blättern Papier der Größe DIN A4 an, indem diese auf drei bzw. vier Spalten mittels senkrechter Trennstriche aufgeteilt wurde. Dabei sind eine vertikale und leicht versetzte Notation sowie Trennstriche nach abgeschlossenen Sinneinheiten erkennbar. TN1 verwendete sowohl Abkürzungen als auch ausgeschriebene Wörter und Symbole, vor allem Pfeile. Die Struktur der Rede ist gut erkennbar.

Die Notizen zur Rede A2 mit Sight-Terp wurden auf zwei Blättern Papier eines Notizblocks der Größe DIN A5 angefertigt, indem das Papier jeweils in der Mitte mit einem Trennstrich auf zwei Spalten aufgeteilt wurde. Auffällig ist hier, dass TN1 vor allem Problemauslöser wie Zahlen, Eigennamen und Termini notierte und dass nach jedem gedolmetschten Segment dieses durchgestrichen wurde. Außerdem wurde die Struktur aufgrund der Nummerierung in Erstens, Zweitens, etc. dank der Ordinalzahlen klar sichtbar gemacht. Im Vergleich zur Rede ohne Tool wurden die Notizen stark reduziert. In beiden Fällen wurden Wörter bzw. Abkürzungen ausschließlich in der Ausgangssprache Englisch notiert.

TN2

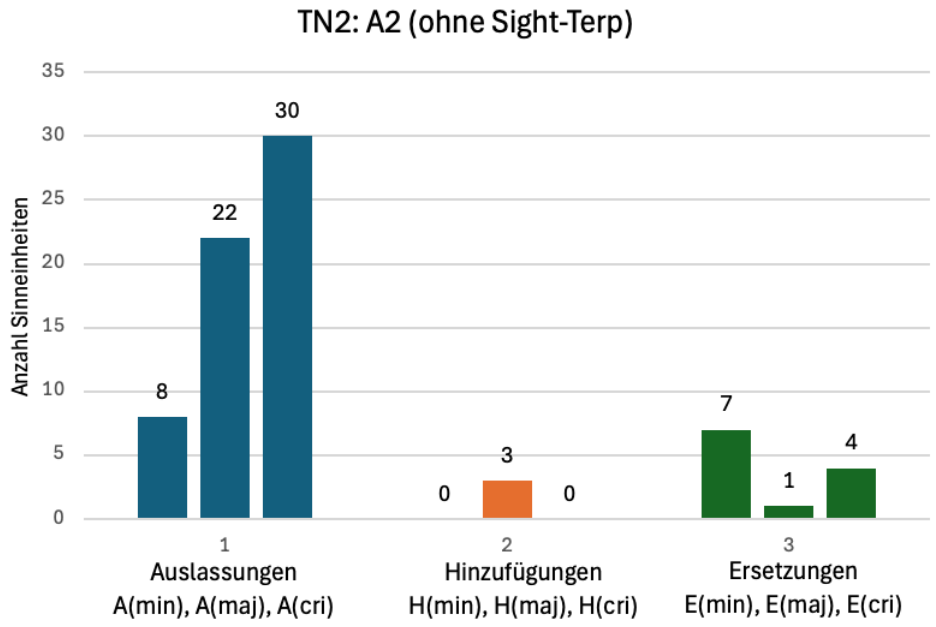


Abbildung 13: Analyseergebnisse von TN2 (ohne Sight-Terp)

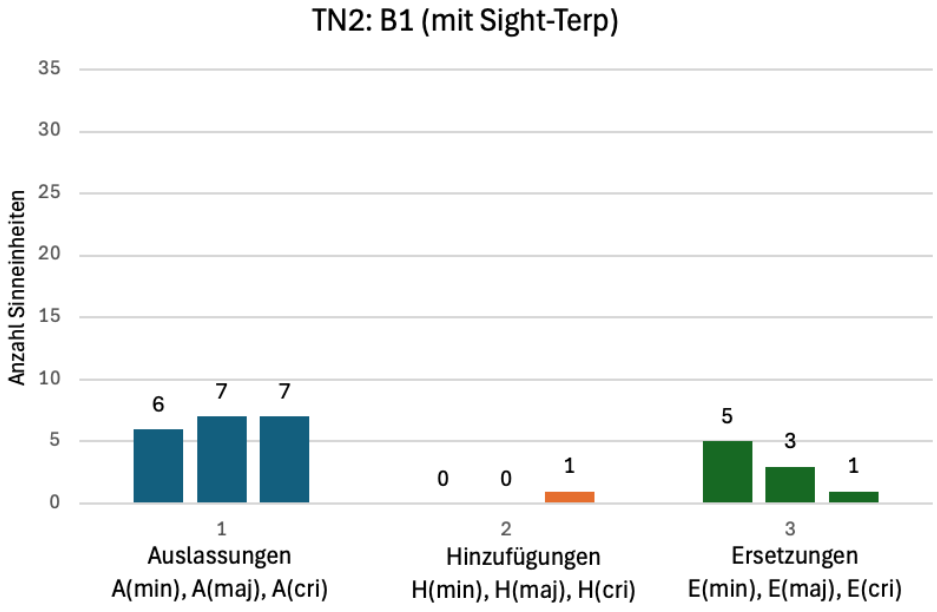


Abbildung 14: Analyseergebnisse von TN2 (mit Sight-Terp)

Die Analyseergebnisse von TN2 zeigen, dass leichte, schwere und kritische Auslassungen grundsätzlich öfter in der Rede A2 vorkamen, welche ohne Sight-Terp gedolmetscht wurde. Es wurden zwei schwere Ersetzungen mehr in der Rede B1 gefunden, ebenso wie eine kritische Hinzufügung, als in der Verdolmetschung von der Rolle der Frauen und der Männer gesprochen wurde, im Original aber der Fokus auf den Frauen lag. Diese Tatsache könnte darauf zurückzuführen sein, dass in der Verdolmetschung mit Sight-Terp der doppelte Input, also jener der Notizen und des Bildschirms, zu einer kurzfristigen kognitiven Überlastung führte. Schwerwiegende Hinzufügungen kamen wiederum nur in der ohne Sight-Terp gedolmetschten Rede A2 vor.

Es zeigt sich auch, dass sieben schwere sowie kritische Auslassungen auch mit Sight-Terp auftraten und somit nicht komplett verhindert werden konnten. Zusätzlich kann zur Rede A2 in der kritischen Fehlerkategorie mit vier Ersetzungen gesagt werden, dass sich die Verdolmetschung ohne Sight-Terp, unabhängig von den Auslassungen, trotzdem stark am Original orientierte. Die Fehlerquote beträgt für die Sinneinheiten in der Rede B1 mit Sight-Terp 18,9 % sowie für die Problemauslöser 18,8 % und für die Sinneinheiten in der Rede A2 ohne Sight-Terp 55,6 % sowie für die Problemauslöser 50,9 %. Diese Ergebnisse zeigen deutlich, dass die Verwendung von Sight-Terp zu einer Verbesserung der Wiedergabe der Sinneinheiten um 36,7 % sowie der Problemauslöser um 36,7 % führte.

TN2 notierte beide Reden auf einem Notizblock der Größe DIN A5 auf je insgesamt sechs Seiten. Beide Reden wurden in der Ausgangssprache Englisch notiert. Bei den angefertigten Notizen wurde jeweils ein Trennstrich in der linken Hälfte gezogen. In der somit entstandenen Spalte wurden u. a. Zahlen, Personalpronomen und Datumsangaben notiert. Die Notationsweise auf der rechten Seite war eher horizontal angelegt. Die Notizen zur Rede A2 ohne Sight-Terp sowie zur Rede B1 mit Sight-Terp ähnelten einander in ihrem Aufbau sowie in der Verwendung von zahlreichen Abkürzungen und Symbolen. TN2 benötigte für beide Reden insgesamt sechs Seiten. Die angefertigten Notizen waren sowohl mit als auch ohne Sight-Terp vergleichbar lang bzw. dicht.

TN3

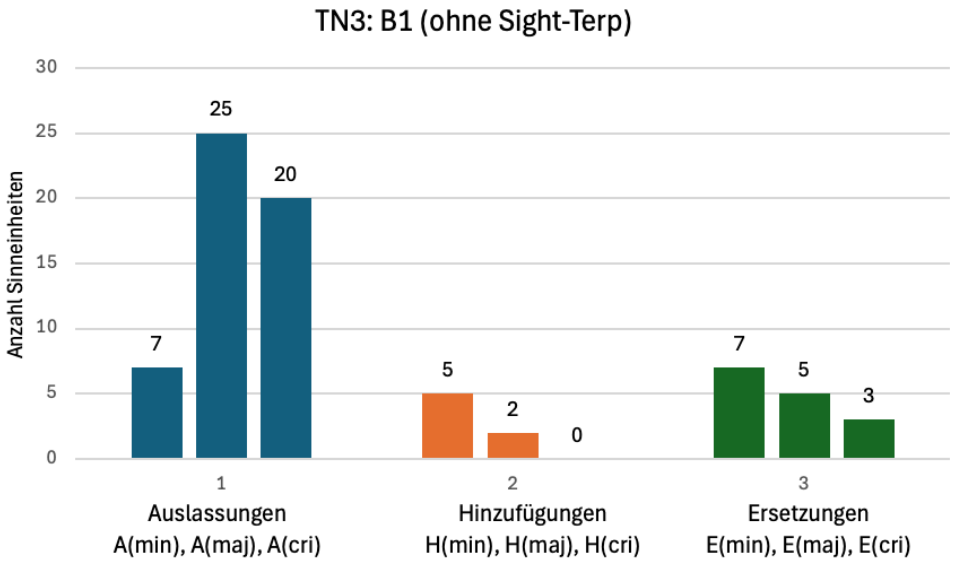


Abbildung 15: Analyseergebnisse von TN3 (ohne Sight-Terp)

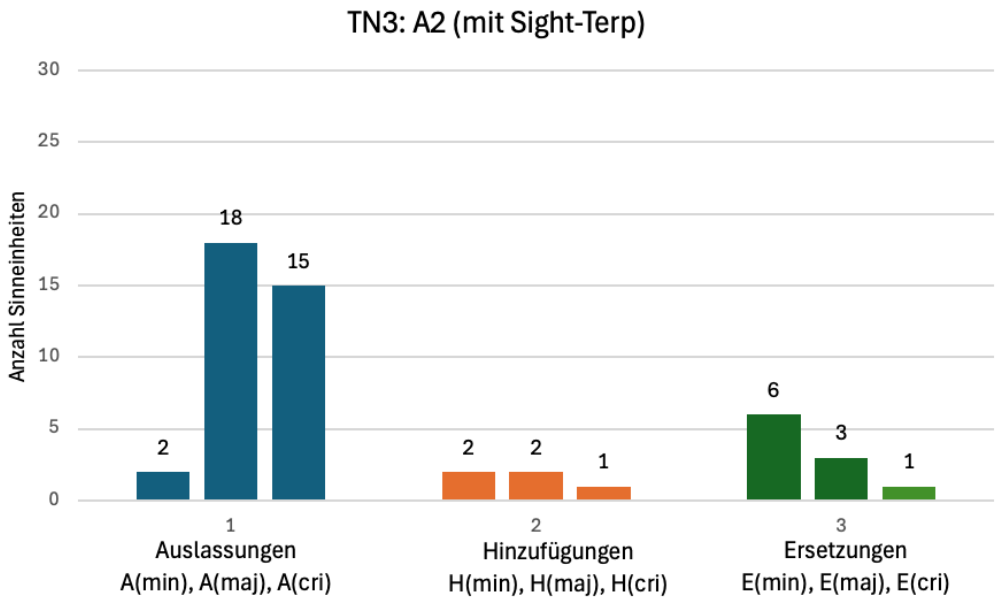


Abbildung 16: Analyseergebnisse von TN3 (mit Sight-Terp)

Die Ergebnisse von TN3 lassen in einigen Fehlerkategorien keinen ganz so starken Unterschied zwischen der Rede ohne Sight-Terp und der Rede mit Sight-Terp erkennen. Dennoch schneidet die Verdolmetschung mit Sight-Terp besser ab. So etwa konnten schwerwiegende Auslassungen um sieben Sinneinheiten in der Verdolmetschung mit der technologischen Unterstützung reduziert werden und kritische Auslassungen um fünf Sinneinheiten. Eine kritische Hinzufügung gab es nur bei der Rede A2, als „das war es auch schon mit dem letzten Punkt“ angefügt wurde, was nicht in der Ausgangsrede gesagt wurde und die Verdolmetschung abrupt und daher möglicherweise auf unerwünscht knappe Weise beendete.

Bei allen anderen Kategorien ermöglichte Sight-Terp quer durch alle Auslassungen, Hinzufügungen und Ersetzungen hindurch eine dem Original adäquatere Verdolmetschung als jene Rede, die ohne die Anwendung gedolmetscht wurde. Die Fehlerquote für TN3 betrug somit für die Rede B1 in Hinblick auf die Sinneinheiten 46,5 % und für die Problemauslöser 33,3 % sowie für die Rede A2 mit Sight-Terp für die Sinneinheiten 37 % und für die Problemauslöser ebenso 33,3 %. Somit konnte TN3 die Verdolmetschung der Sinneinheiten dank Sight-Terp um 9,5 % verbessern, die korrekte Verdolmetschung der Problemauslöser blieb hingegen für die Rede sowohl mit als auch ohne Sight-Terp konstant.

Die Notizen von TN3 wurden auf einem Notizblock der Größe DIN A5 angefertigt. Bei beiden Durchgängen wurden keine Trennstriche gemacht. Die Sprache war in beiden Fällen Englisch. Beide Reden wurden eher horizontal notiert. Teilweise wurden Aufzählungen eingerückt. Grundsätzlich konnte jedoch keine systematisch angewandte Methode erkannt werden. Es wurden mehr Wörter und Abkürzungen als Symbole verwendet. Die Menge an notierten Elementen war sowohl bei der Rede mit als auch ohne Sight-Terp vergleichbar. Insgesamt wurden für beide Reden vier Blätter Papier benötigt.

TN4

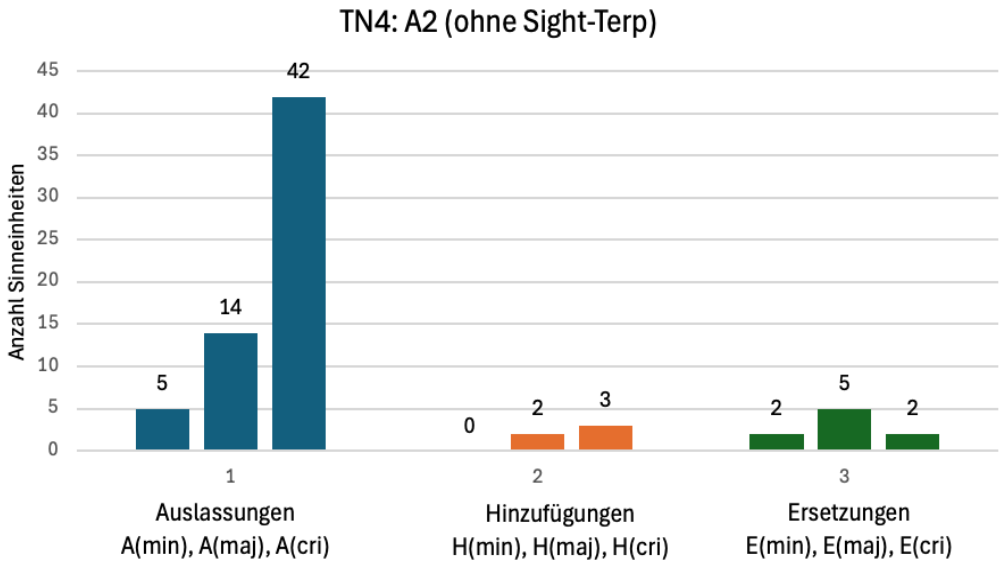


Abbildung 17: Analyseergebnisse von TN4 (ohne Sight-Terp)

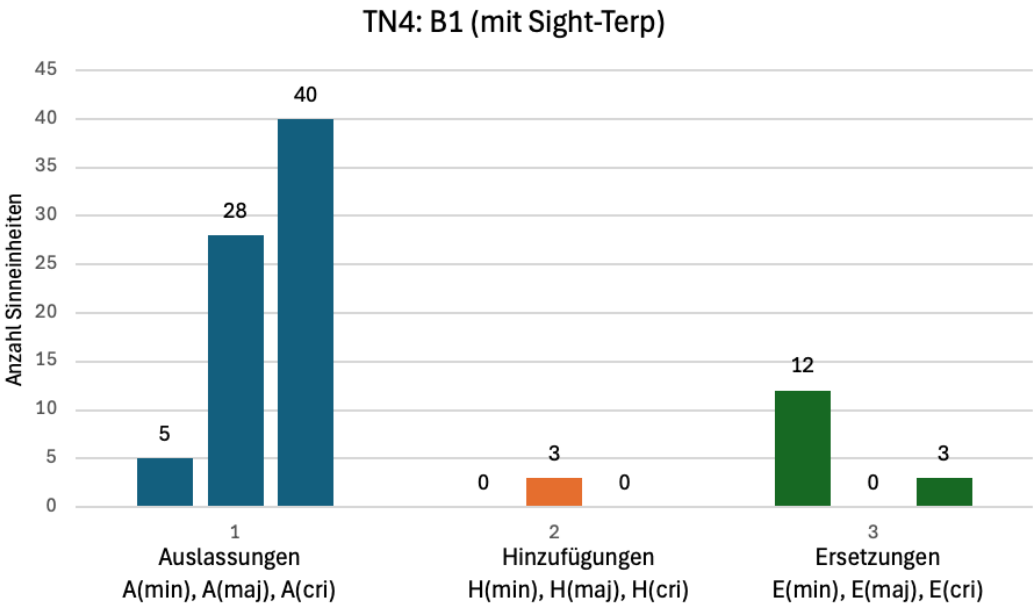


Abbildung 18: Analyseergebnisse von TN4 (mit Sight-Terp)

TN4 verdolmetschte die Rede A2 ohne Sight-Terp und die Rede B1 mit Sight-Terp. Auffällig ist hier, dass in der Verdolmetschung der Rede B1 mit technologischer Unterstützung mehr leichte Ersetzungen, ebenso wie schwere Auslassungen und Hinzufügungen, festgestellt wurden. Bei den kritischen Auslassungen wurden um zwei Sinneinheiten weniger ausgelassen, also es wurde etwas genauer im Durchgang mit Sight-Terp gedolmetscht. Im Vergleich zu den anderen Teilnehmenden ist dieses Ergebnis auffällig, da relativ viele Auslassungen in beiden Reden vorkamen. Eine mögliche Erklärung ist, dass die Verwendung von Sight-Terp mit kognitiven Belastungen (vgl. Abschnitt 1.2.1) einhergeht, welche aufgrund der erlernten und erprobten Notizentechnik weniger stark zum Vorschein kommen.

Letztlich wies TN4 bei der Rede A2 ohne Sight-Terp eine Fehlerquote bei den Sinneinheiten von 55,6 % sowie bei den Problemauslösern von 50,9 % auf und für die Rede B1 mit Sight-Terp betrugen diese Werte jeweils 57,2 % und 52,1 %. TN4 verschlechterte sich also leicht unter Anwendung von Sight-Terp jeweils um 1,7 % bei den Sinneinheiten und um 1,2 % bei den Problemauslösern.

TN4 verwendete für die Verdolmetschung beider Reden einen Notizblock der Größe DIN A5. Es wurde jeweils in der linken Hälfte ein vertikaler Trennstrich gezogen, jedoch wurde in die dabei entstandene linke Spalte kaum etwas eingetragen. Beide Reden wurden in der Ausgangssprache notiert und für beide Reden benötigte TN4 drei Seiten. Auffällig ist hier, dass sowohl mit als auch ohne Sight-Terp Wörter zumeist ausgeschrieben wurden. In der Rede ohne Sight-Terp war infolge der strukturierten Ausgangsrede auch eine Chronologie in den Notizen dank der Nummerierung in Erstens, Zweitens, etc. erkennbar. Ansonsten konnten keine Unterschiede zwischen den Durchgängen ausgemacht werden.

TN5

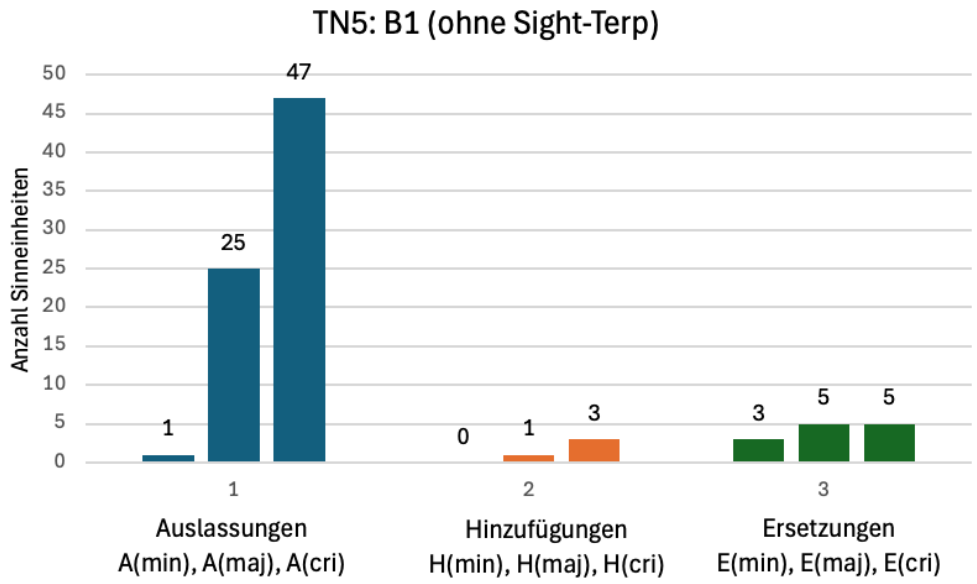


Abbildung 19: Analyseergebnisse von TN5 (ohne Sight-Terp)

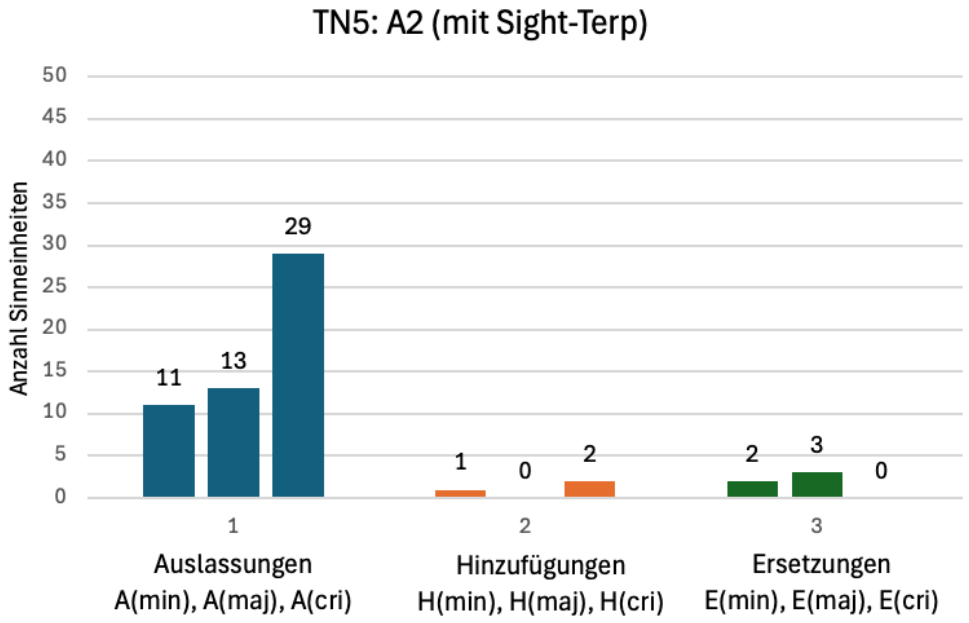


Abbildung 20: Analyseergebnisse von TN5 (mit Sight-Terp)

Die Abbildungen 19 und 20 zeigen die fehlerhaften Sinneinheiten der durch TN5 verdolmetschten Rede A2 mit Sight-Terp und B1 ohne Sight-Terp. Dabei zeigt sich, dass insbesondere leichte Auslassungen öfter unter Verwendung des CAI-Tools auftauchten. Bei allen anderen Kategorien zeigen die Ergebnisse mit Sight-Terp eine Verbesserung in der Verdolmetschung. Beispielweise konnten schwere und kritische Auslassungen mit Sight-Terp in der Rede A2 deutlich reduziert werden. Interessant ist, dass TN5 bei der Verdolmetschung ohne Sight-Terp fünf kritische Ersetzungen einbaute. Diese betrafen in allen fünf Fällen die fehlerhafte Verdolmetschung der wahrscheinlich herausforderndsten Problemauslöser Zahlen (vgl. dazu Seeber, 2015: 85f.). Es wurde statt „one in three women" „eine von fünf Frauen" gedolmetscht sowie statt „after the age of 15" „bis zum Alter von 15" und statt „more than a half" einfach nur „mehr". Außerdem wurden aus „one in twenty" „eine von zehn" und letztlich wurde statt „75 %" „eine von 5" gesagt. Hier zeigt sich deutlich, dass die ausgewählten Qualitätsaspekte (vgl. Kapitel 2) in der Verdolmetschung ohne Sight-Terp negativ beeinträchtigt wurden.

Die Fehlerquote beträgt für TN5 bei der Rede B1 ohne Sight-Terp für die Sinneinheiten 56,6 % und für die Problemauslöser 54,2 % sowie für die Rede A2 mit Sight-Terp für die Sinneinheiten 45,2 % und für die Problemauslöser 43,9 %. Sight-Terp erlaubte also eine Verbesserung der gedolmetschten Sinneinheiten um 11,4 % und jene der Problemauslöser um 10,3 %.

TN5 fertigte die Notizen auf einem Notizblock der Größe DIN A5 an und zog zwei Trennstriche. Es wurde in beiden Fällen vor allem in der zweiten und dritten Spalte und horizontal notiert. Zudem verwendete TN5 sehr wenige Abkürzungen bzw. Symbole und notierte in beiden Durchgängen auf Englisch. Für beide Reden wurden drei Seiten benötigt. Es konnten keine weiteren Auffälligkeiten bei den angefertigten Notizen im Vergleich zwischen der Verdolmetschung mit bzw. ohne Sight-Terp erkannt werden.

TN6

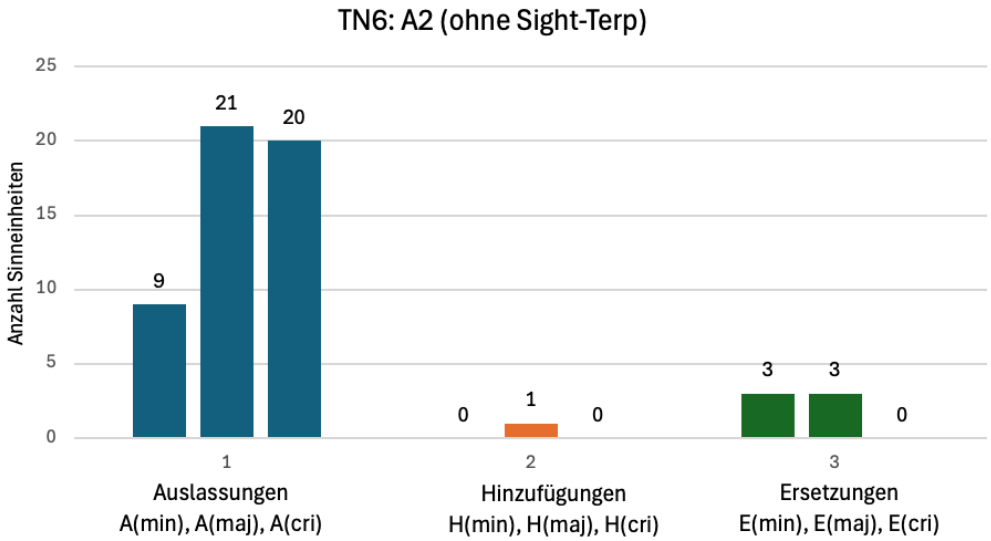


Abbildung 21: Analyseergebnisse von TN6 (ohne Sight-Terp)

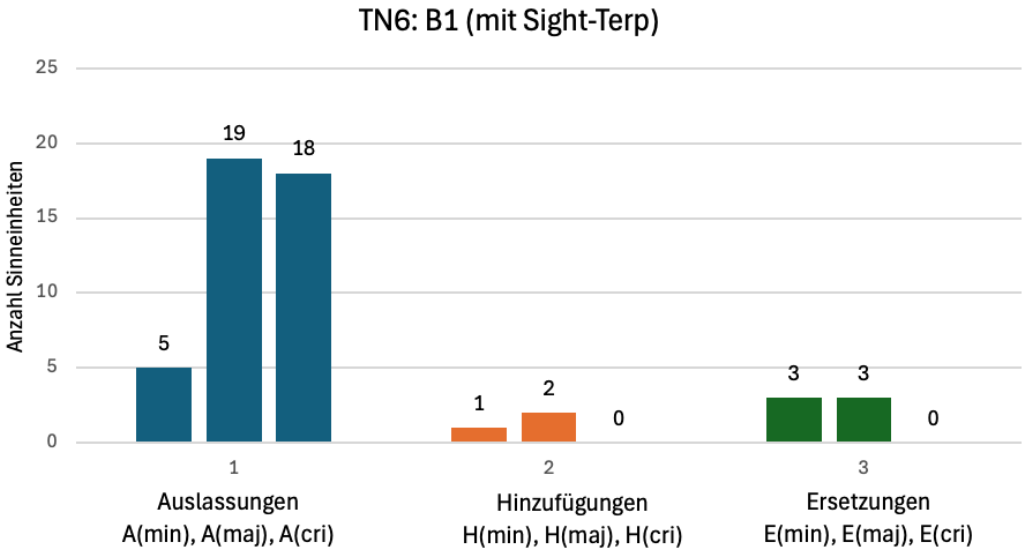


Abbildung 22: Analyseergebnisse von TN6 (mit Sight-Terp)

TN6 verdolmetschte die Rede A2 ohne Sight-Terp und B1 mit Sight-Terp. Die Abbildungen 21 und 22 lassen auf einen Blick erkennen, dass die Verwendung von Sight-Terp auch hier zu einer Verbesserung bzw. zumindest zu einem konstanten Niveau an Auslassungen, Hinzufügungen und Ersetzungen aller Kategorien beitragen konnte. Die Unterschiede zwischen den jeweiligen Kategorien sind zwar nicht sehr groß, jedoch erlaubte Sight-Terp eine leicht vollständigere und genauere Verdolmetschung der herangezogenen Rede. Hervorzuheben ist, dass bei den Verdolmetschungen von TN6 keine kritischen Hinzufügungen bzw. Ersetzungen auftraten, weder in der Rede A2 noch in der Rede B1. Dieses Ergebnis lässt eine durchwegs genaue Arbeitsweise vermuten.

Jedoch zeigt sich, dass bei der Rede A2, die ohne Sight-Terp gedolmetscht wurde, bedeutungstragende Problemauslöser häufig ausgelassen wurden und als kritische Auslassungen von dementsprechenden Sinneinheiten ebenso in die Analyse einflossen. Die Fehlerquote von TN 6 beträgt für die Rede A2 ohne Sight-Terp in Bezug auf die Sinneinheiten 42,2 % und für die Problemauslöser 49,1 % sowie für die Rede B1 mit Sight-Terp für die Sinneinheiten 32,1 % und für die Problemauslöser 25 %. Dank Sight-Terp konnte sich TN6 in der Verdolmetschung mit Sight-Terp um 10,1 % bei den Sinneinheiten und um 24,1 % bei den Problemauslösern verbessern.

TN6 notierte auf einem DIN A5-Notizblock, indem ein vertikaler Trennstrich in der linken Hälfte gezogen wurde. Beide Reden wurden auf Englisch auf fünf Seiten notiert. Teilweise wurden zwei Gedankenelemente mit Strichen verknüpft. Es konnten vor allem ausgeschriebene und abgekürzte Wörter erkannt werden. Zudem kann trotz einer horizontal orientierten Notationsweise gesagt werden, dass oft großzügige Abstände zwischen den Notizen zu einem Überblick führen konnten. Es konnten keine auffälligen Unterschiede in den beiden Durchgängen festgestellt werden.

TN7

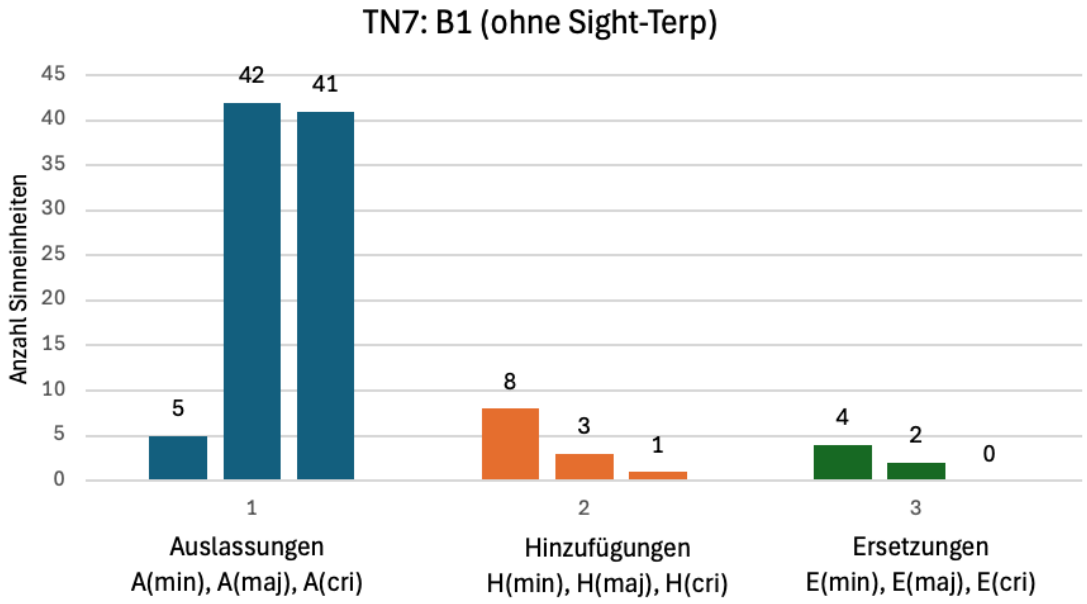


Abbildung 23: Analyseergebnisse von TN7 (ohne Sight-Terp)

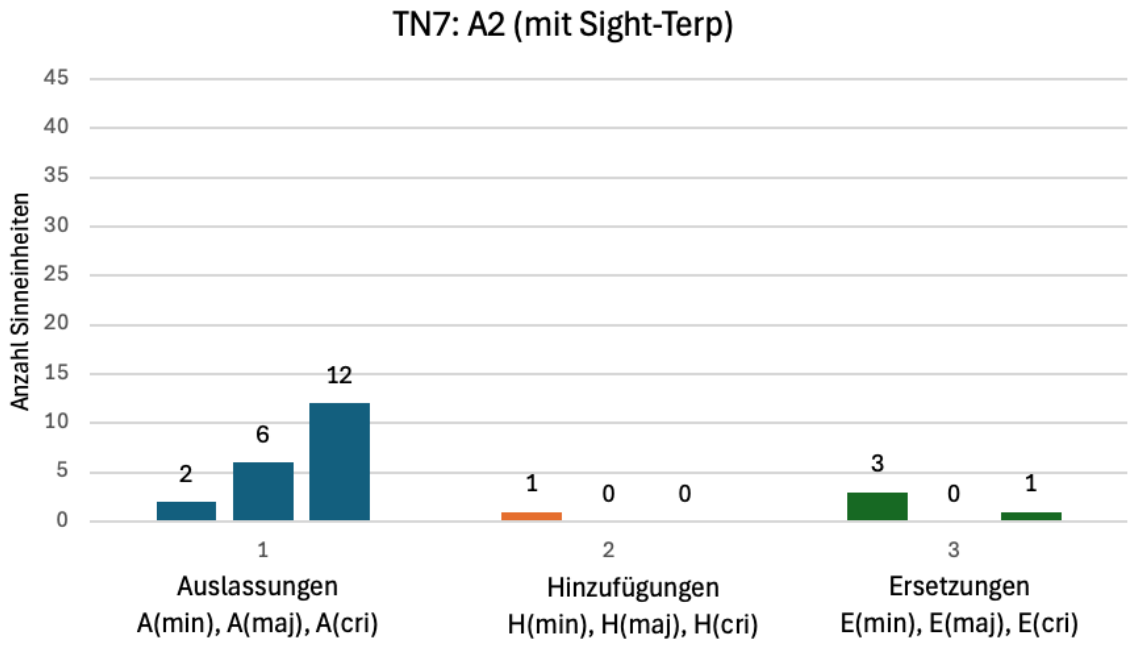


Abbildung 24: Analyseergebnisse von TN7 (mit Sight-Terp)

TN7 verdolmetschte die Rede A2 mit Sight-Terp und die Rede B1 ohne Sight-Terp. Die Abbildungen 23 und 24 zeigen eindeutig, dass alle Fehlerkategorien der Sinneinheiten bei der Verdolmetschung mit Sight-Terp deutlich reduziert werden konnte. Besonders sichtbar ist dies wieder anhand der Vollständigkeit, denn schwere und kritische Auslassungen kamen sehr stark in der Rede B1, die ohne Sight-Terp bzw. nur mit Notizen gedolmetscht wurde, vor. Die kritischen Auslassungen betrafen in der Rede B1 nicht zuletzt bedeutungstragende Problemauslöser, neben Zahlen zu einem Großteil auch Termini. Daneben waren ebenso weitere Sinneinheiten negativ betroffen, die zu einem übergreifenden Verständnis bzw. einer allgemeinen Sinnübereinstimmung beitragen sollten. Jedoch kam es auch in der Verdolmetschung mit Sight-Terp zu einigen fehlerhaften Sinneinheiten bzw. Fehlern bei sinntragenden Problemauslösern, welche in erster Linie die Vollständigkeit beeinträchtigten bzw. in einem Fall aufgrund einer Ersetzung einen Sinn hinzufügten, der die ursprüngliche Aussage änderte. Hier wurde statt „office buildings“ „öffentliche Einrichtungen“ gedolmetscht.

Die Fehlerquote beträgt für TN7 für die Rede B1 ohne Sight-Terp in Hinblick auf die Sinneinheiten 66,7 % sowie für die Problemauslöser 37,5 % und für die Rede A2 mit Sight-Terp kommen diese Werte auf 18,5 % für die Sinneinheiten und 22,8 % für die Problemauslöser. TN7 konnte sich beim Dolmetschen mit der CAI-Anwendung um 48,2 % bei der korrekten Wiedergabe der Sinneinheiten sowie um 14,7 % bei der Wiedergabe der Problemauslöser verbessern.

Die Notizen von TN7 wurden im Format DIN A5 angefertigt und für beide Durchgänge, mit und ohne CAI-Tool, wurden insgesamt vier Seiten benötigt. Es wurden in beiden Fällen Zahlen, Symbole, Abkürzungen und ausgeschriebene Wörter verwendet und die Sprache der Notizen war Englisch. Die Notizen wurden durch einen linksorientierten Trennstrich aufgeteilt, in der linken Spalte wurde jedoch kaum notiert. Außerdem wurde eine leichte Vertikalität erkannt. Es konnten keine auffälligen Unterschiede in der Notationsweise ohne bzw. mit Sight-Terp festgestellt werden.

TN8

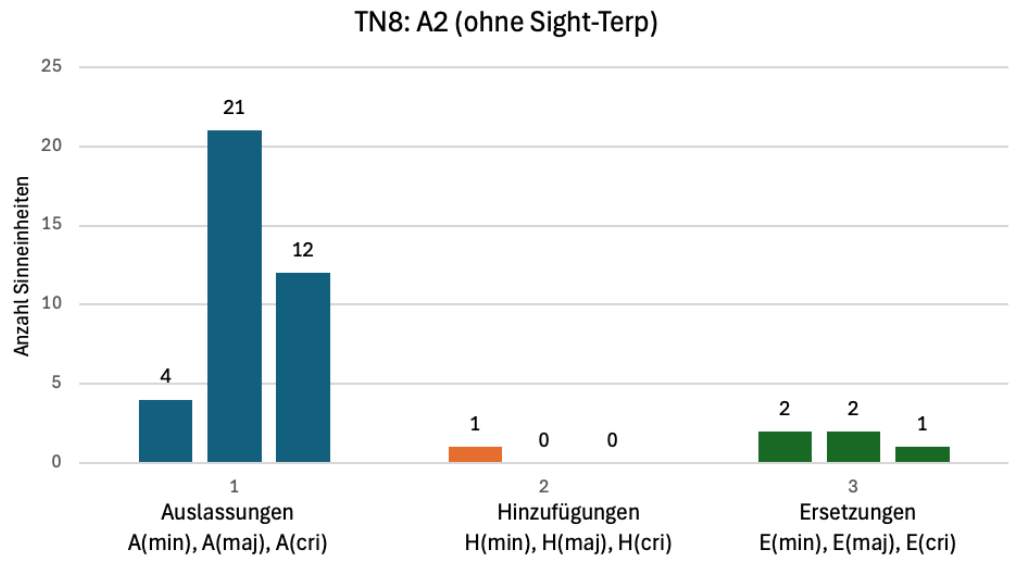


Abbildung 25: Analyseergebnisse von TN8 (ohne Sight-Terp)

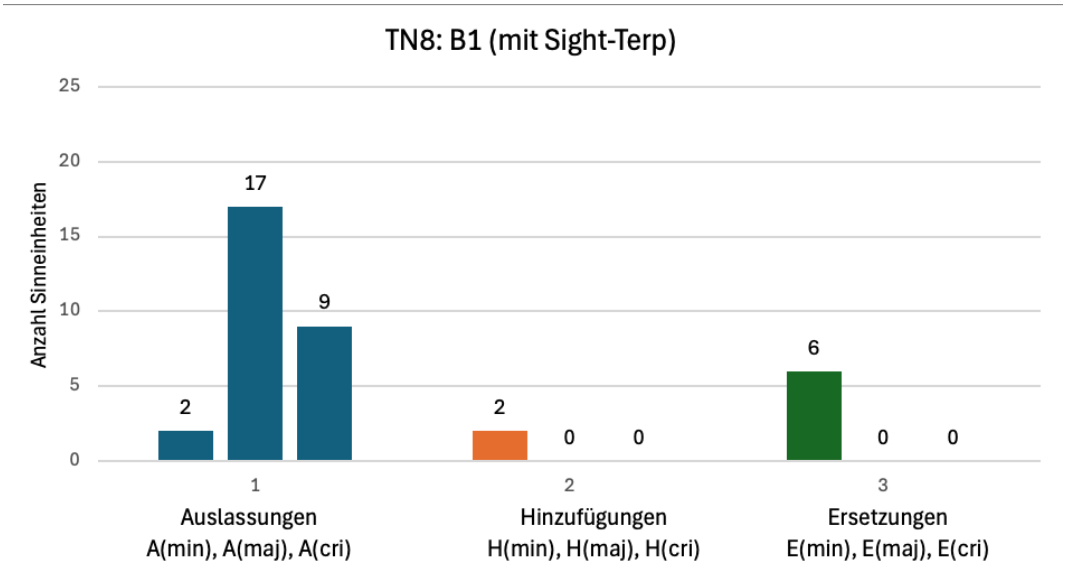


Abbildung 26: Analyseergebnisse von TN8 (mit Sight-Terp)

Die beiden Abbildungen 25 und 26 zeigen, dass bei der Verdolmetschung der Rede A2 ohne Sight-Terp nur etwas mehr Fehler erkannt werden konnten als bei der gedolmetschten Rede B1, welche mit Sight-Terp erfolgte. Zudem gab es in der Rede mit Sight-Terp mehr leichte Ersetzungen. Ansonsten kann festgehalten werden, dass alle Abstufungen der Fehlerkategorie Auslassungen dank Sight-Terp reduziert werden konnten. Besonders hervorzuheben ist auch, dass TN8 (fast) keine schweren sowie kritischen Hinzufügungen bzw. Ersetzungen bei den beiden Reden einbaute, lediglich solche leichter Art. Dieses Ergebnis könnte auf die Tatsache zurückzuführen sein, dass TN8, so wie alle anderen Teilnehmenden auch, im Studium bereits fortgeschritten ist (vgl. Abschnitt 4.1). Denn auch die kritischen Auslassungen traten bei beiden Reden eher selten auf.

Somit führten diese Ergebnisse zu einer Fehlerquote bei der Rede A2 ohne Sight-Terp für die Sinneinheiten von 31,9 % sowie für die Problemauslöser von 31,6 % und für die Rede B1 mit Sight-Terp ergab dies eine Fehlerquote von 22,6 % bei den Sinneinheiten sowie von 16,7 % bei den Problemauslösern. TN8 verbesserte sich mit Sight-Terp um 9,2 % bei den Sinneinheiten und um 14,9 % bei den Problemauslösern.

TN8 notierte auf einem Notizblock der Größe DIN A5 und benötigte für die Rede A2 ohne Sight-Terp sieben Blätter, für die Rede B1 mit Sight-Terp fünf Blätter. Bei beiden Reden wurde keine vertikale Trennlinie gezogen, jedoch wurden kurze horizontale Striche nach Gedankenelementen gemacht. Die Notizen zu beiden Reden wiesen zahlreiche Abkürzungen, Symbole und eine eher vertikale bzw. lockere Notation auf. Die Sprache war in beiden Fällen Englisch. Der einzige hier festgestellte Unterschied ist, dass in der Rede B1 mit Sight-Terp etwas weniger notiert wurde.

TN9

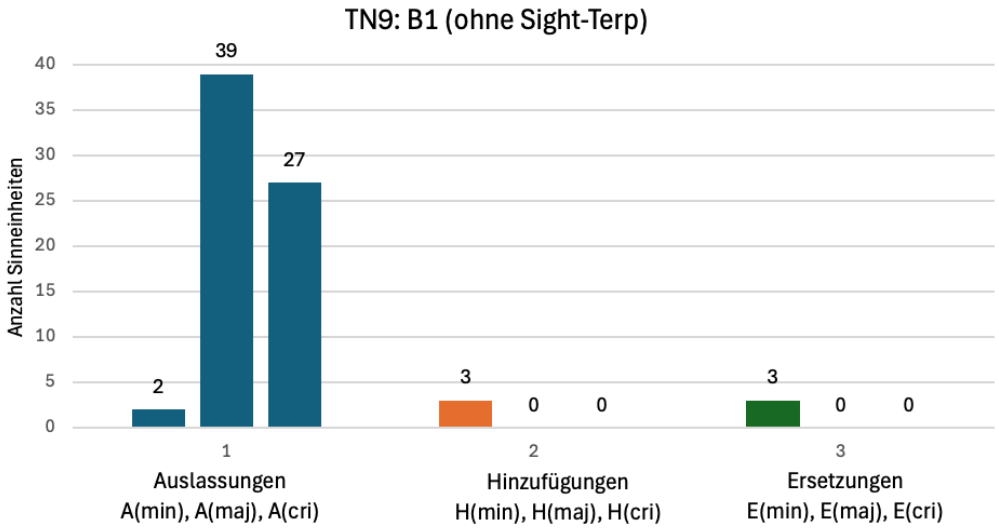


Abbildung 27: Analyseergebnisse von TN9 (ohne Sight-Terp)

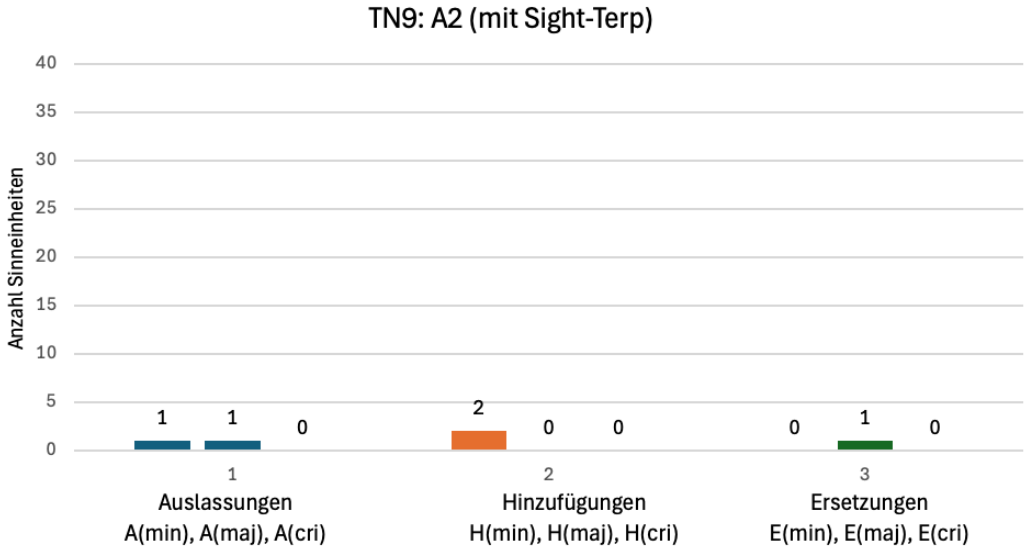


Abbildung 28: Analyseergebnisse von TN9 (mit Sight-Terp)

TN9 verdolmetschte die Rede A2 mit Sight-Terp und die Rede B1 ohne Sight-Terp. Die Abbildungen 27 und 28 veranschaulichen, dass Sight-Terp nicht nur die kritischen Fehlerkategorien komplett ausschließen konnte, es wurden auch die schweren bzw. leichten Fehler auf ein Minimum beschränkt. Unter alleiniger Verwendung der Notizentechnik zeigte sich, dass insbesondere wesentlich mehr schwere, aber auch kritische Auslassungen bei der Rede B1 über Gewalt gegen Frauen auftauchten. Eine Analyse der gemachten Notizen zeigte, dass sich TN9 bei der Verwendung von Sight-Terp fast keine Notizen machte. Daher kann davon ausgegangen werden, dass für die Verdolmetschung hauptsächlich bzw. lediglich Sight-Terp herangezogen wurde und offensichtlich zu einem sehr guten Ergebnis führte.

Die Fehlerquote von TN9 für die Rede B1 ohne Sight-Terp beträgt für den Aspekt der Sinneinheiten insgesamt 46,5 %, sowie für die Problemauslöser 35,4 % und für die Rede A2 mit Sight-Terp lediglich 3,7 % für die Sinneinheiten und 1,8 % für die Problemauslöser. Sight-Terp ermöglichte im vorliegenden Fall eine sehr starke Verbesserung der Verdolmetschung der Sinneinheiten um 42,8 % und jene der Problemauslöser um 33,7 %.

Die von TN9 angefertigten Notizen befanden sich auf einem DIN A5-Block, bei dem eine vertikale Trennlinie in der linken Hälfte gezogen wurde. Im Zuge der Verdolmetschung der Rede A2, bei der Sight-Terp verwendet wurde, fertigte TN9 fast keine Notizen an, es wurden lediglich zwei Elemente, nämlich „Wasserspeichertunnel“ und „3 Mrd“ notiert. Bei der Rede B1, die ohne Sight-Terp gedolmetscht wurde, wurden ein paar Symbole, Zahlen, Abkürzungen und Wörter, ausschließlich auf Englisch, erkannt. Die Notation erfolgte horizontal und teilweise versetzt sowie auch aufgelockert. TN9 benötigte für die Verdolmetschung ohne die technologische Anwendung fünf Seiten, für die Verdolmetschung mit Sight-Terp wurde fast nichts notiert, bis auf die oben erwähnten zwei Elemente auf knapp einer halben Seite. Somit wurden die Notizen mit Sight-Terp stark reduziert bzw. fast gar nicht gemacht.

TN10

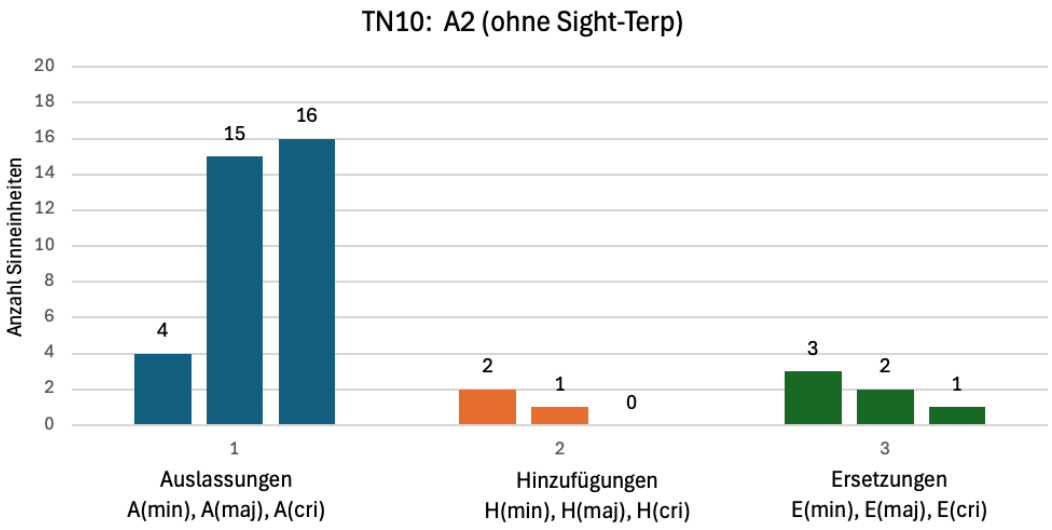


Abbildung 29: Analyseergebnisse von TN10 (ohne Sight-Terp)

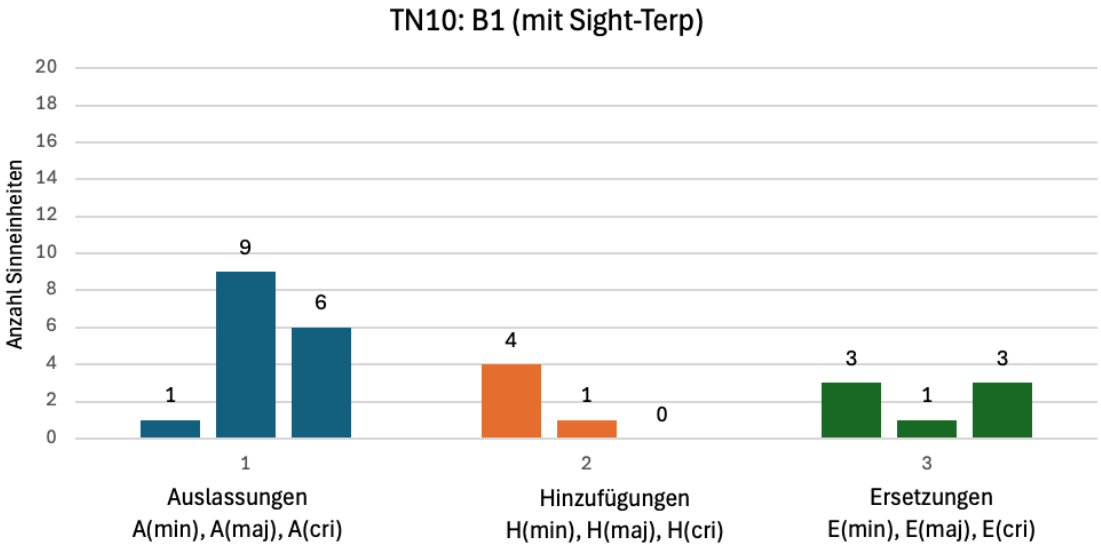


Abbildung 30: Analyseergebnisse von TN10 (mit Sight-Terp)

Die Abbildungen 29 und 30 veranschaulichen die fehlerhaften Sinneinheiten von TN10 für die Rede B1, die mit Sight-Terp gedolmetscht wurde, und jene für die Rede A2, die ohne Verwendung von Sight-Terp gedolmetscht wurde. Dabei zeigt sich, dass die Rede B1 mit Sight-Terp in der Verdolmetschung der Sinneinheiten grundsätzlich besser abschnitt. Jedoch gab es hierbei mit drei kritischen Ersetzungen eine etwas erhöhte Anzahl dieser Fehlerkategorie im Vergleich zur gedolmetschten Rede ohne Sight-Terp. Dies betraf einerseits die Verdolmetschung von einem bedeutungstragenden Problemauslöser, nämlich „after the age of 15“ als „nach dem 5. Lebensjahr“, sowie auch von zwei weiteren Sinneinheiten, und zwar „another unconnected factor“ als „ein weiterer zusammenhängender Faktor“ und von „depressing reality“ als „definierende Realität“.

Insgesamt gibt vor allem die kritische Fehlerkategorie zu erkennen, dass die Verdolmetschung mit Sight-Terp das Ergebnis positiv beeinflussen konnte. Die Fehlerquote von TN10 für die Rede A2 ohne Sight-Terp beträgt für die Sinneinheiten 32,6 % und für die Problemauslöser 31,6 % sowie für die Rede B1 mit Sight-Terp für die Sinneinheiten 17,6 % und für die Rede A2 ohne Sight-Terp 12,5 %. Somit führte Sight-Terp auch hier zu einer Verbesserung zwischen der Verdolmetschung ohne bzw. mit Sight-Terp von knapp 15 % bei den Sinneinheiten sowie von 19,1 % für die Problemauslöser.

TN10 notierte ebenso auf einem Block der Größe DIN A5 und zog einen links orientierten Trennstrich. In der entstandenen linken Spalte wurden Schlagwörter bzw. Ordinalzahlen notiert. Die Notationsweise war in beiden Durchgängen eher horizontal und die Sprache war die Ausgangssprache Englisch. Es konnten einige Abkürzungen, Symbole und Zahlen sowie auch einige ausgeschriebene Wörter erkannt werden. Für beide Reden wurden zwei Seiten benötigt. Es konnten keine auffälligen Unterschiede in der Notizenanfertigung mit bzw. ohne Sight-Terp festgestellt werden.

Alle TN: Dauer der Verdolmetschungen mit/ohne Sight-Terp

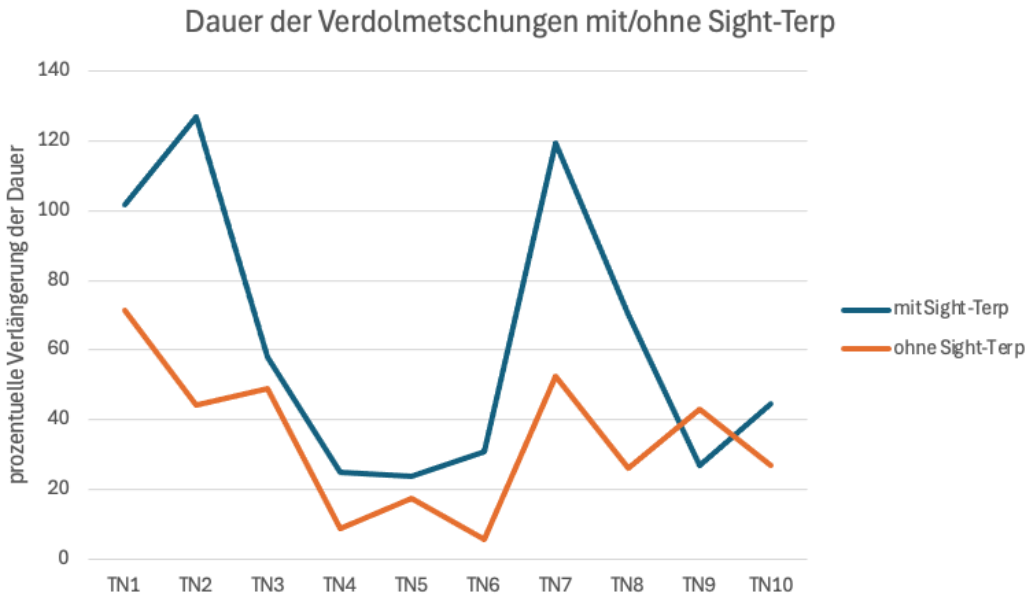


Abbildung 31: Dauer der Verdolmetschungen mit/ohne Sight-Terp

Abbildung 31 zeigt eine Gegenüberstellung der prozentuellen Verlängerung der Dauer der durch die Teilnehmenden jeweils verdolmetschten Reden A2 bzw. B1 mit bzw. ohne Sight-Terp im Vergleich zur Dauer des ursprünglichen Redebeitrags in Minuten und Sekunden.

Dabei sind Unterschiede bei der Dauer der Verdolmetschung mit bzw. ohne CAI-Tool zu erkennen. Alle Teilnehmenden bis auf TN9 benötigten für die konsekutive Verdolmetschung unter der Versuchsbedingung mit Sight-Terp grundsätzlich mehr Zeit als das Original. Besonders auffällig sind die um ein Vielfaches verlängerten Verdolmetschungen von TN1, TN2 und TN7 unter Verwendung von Sight-Terp.

Aus diesen Ergebnissen lässt sich grundsätzlich ableiten, dass die Verwendung von Sight-Terp die Konsekutivverdolmetschung verlängert. Mögliche Gründe könnten weitere visuelle Inputs sein, nämlich die Transkription und die maschinelle Übersetzung, die zum Dolmetschen und den möglicherweise selbst angefertigten Notizen hinzukommen und den Koordinationsaufwand erhöhen.

Im Folgenden Abschnitt erfolgt die Darstellung der Analyseergebnisse aus den Antworten der Teilnehmenden beim durchgeführten Leitfaden-Interview. Diese werden für die gestellten Fragen anhand einer Kategorieneinteilung präsentiert.

5.2 Leitfaden-Interview

Kategorie 1: Schwierigkeit der Ausgangsreden

Auf die erste Frage antworteten die Teilnehmenden sehr einheitlich. Die Teilnehmenden waren grundsätzlich auf beide Reden thematisch vorbereitet. Die größte Herausforderung betraf jedoch ein sehr hohes Redetempo der Ausgangsreden. Dementsprechend wurde das Notieren insbesondere im Durchgang ohne Sight-Terp beeinflusst: „(…) die Rede und meine Notizen (waren) beide dicht“ (TN5, Pos. 13).

Außerdem gaben die Teilnehmenden einen Unterschied im Schwierigkeitsgrad an. Den Aussagen der Teilnehmenden zufolge war die Rede über Erdbeben in Japan aufgrund der klaren Einteilung in acht Maßnahmen wesentlich übersichtlicher und strukturierter als die Rede über Gewalt gegen Frauen und erlaubte es somit, bereits beim Notieren eine Übersicht zu schaffen. Der Schwierigkeitsgrad der Rede über Gewalt gegen Frauen wurde zudem neben einer schnellen Vortragsgeschwindigkeit, die für beide Reden vergleichbar war (vgl. TN1, TN4, TN9, TN10), aufgrund einer hohen Informationsdichte bzw. Anzahl an schnell hintereinander auftauchenden Problemauslösern wie Zahlen, Eigennamen, etc. weiter verstärkt.

Kategorie 2: Allgemeine Arbeitsweise mit Sight-Terp

Bei den Antworten in dieser Kategorie fielen starke Unterschiede in der Einstellung der Teilnehmenden gegenüber dem CAI-Tool Sight-Terp auf. Die Aussagen zur Arbeitsweise und zu einem möglichen Einsatz der Anwendung spiegeln diese wider.

Zunächst äußerten sich die TN oft sehr begeistert, etwa TN1 (Pos. 21): „(…) das erste Mal, dass ich eine Software gefunden habe, die nicht eine Ablenkung ist, sondern wirklich ein Helping Tool (…)“, oder auch TN2 (Pos. 5). „Das war doch auf jeden Fall eine große Hilfestellung (…)“. TN9 (Pos. 29) meinte sogar: „Es ist sehr einfach zu verstehen (…), ich habe dem blind vertraut (…)“.

Nach weiteren Überlegungen in Hinblick auf eine Verwendung der Anwendung wurden die Stimmen tendenziell etwas kritischer. TN3 (Pos. 9) sah im Gegensatz zu TN9 den größten Nachteil genau darin, dass „man sich auf das Tool verlässt“. Zudem führten einerseits eine fehlende Struktur bzw. verschobene Absätze und fehlende Markierungen in Sight-Terp zu weiteren abgeneigten Einstellungen (vgl. TN1, TN2, TN9, TN10). Außerdem wurde erwähnt, dass der Text dazu verleiten würde, dass man stärker „an den Wörtern als an dem Inhalt“ hängt (vgl. TN8, Pos. 18). Diese Überlegung geht auf die im Abschnitt 1.2.2 zur

Notizentechnik angesprochenen Aspekte des Notierens von Bedeutungen und zumeist Symbolen, Abkürzungen, etc. ein, welche eine „Abkehr vom Wort als Worthülse“ bedeuten (vgl. Matyssek, 2012: 37f.).

Auf die Frage, ob die Teilnehmenden Sight-Terp weiterverwenden würden, fielen die Antworten teils negativ aus. TN3 (Pos. 17) behauptete: „Es hat mir (…) bei den Stellen, die ich benötigt habe, geholfen, aber ich würde Sight-Terp nicht verwenden, weil ich (…) mich sicherer fühle, wenn ich meine Notizen habe, die ich selbst notiert habe mit dem, was ich verstanden habe“. TN4 (Pos. 25) erzählte nuancierter, dass die Verwendung von Sight-Terp im Grunde vorstellbar wäre, aber „nur wenn es kürzere Abschnitte sind (…)“.

Zur Arbeitsweise wurde außerdem kritisch angemerkt, dass unter Zuhilfenahme von Sight-Terp beim Konsekutivdolmetschen der dolmetschenden Person drei Texte, nämlich die transkribierte Ausgangsrede, die maschinell übersetzte Version und die angefertigten Notizen, zur Verfügung stehen würden, die während des Dolmetschens potenziell Verwirrung stiften (vgl. TN5, Pos. 13).

Aus den Antworten ließ sich weiter ableiten, dass die Teilnehmenden gegenüber der Idee einer technologischen Unterstützung beim Konsekutivdolmetschen grundsätzlich eher positiv eingestellt waren, dass sie sich über die tatsächliche Anwendung in ihrer aktuellen Form jedoch (noch) nicht ganz im Klaren waren.

Möglicherweise könnte neben einer vermehrten Übung mit dem Tool (vgl. TN1, TN3, TN9) dieses schlichtweg als Ergänzung zur Notizentechnik im Falle fehlender Informationen betrachtet werden und so eingesetzt werden. TN6 (Pos. 17) fügte hinzu: „Das war schon beruhigend, um nochmal eine Gegenprüfung zu haben“, oder „(…), wenn man den roten Faden wieder aufnehmen möchte (…)“ (TN2, Pos. 21).

Eine interessante Überlegung zur Anwendung von Sight-Terp in einem realen Dolmetscheinsatz kam von TN2 (Pos. 25), da hier angesprochen wurde, dass es für die Kunden und Kundinnen „(…) ein komisches Bild macht, wenn sie sehen, da läuft was mit (…)“. TN9 (Pos. 49) hinterfragte ebenso: „(…) ich glaube sie hätten es in dem Fall komisch gefunden, wenn ich da sitze mit meinem Laptop (…)“. Diese Befürchtung einer Infragestellung der intellektuellen Arbeit des Dolmetschens bei der Verwendung von prozessorientierten Technologien wurde auch von Fantinuoli (2018a: 163) angesprochen.

Möglicherweise ist diese Befürchtung auf die (noch) nicht ganz klare Arbeitsweise mit der Anwendung zurückzuführen. TN10 (Pos. 33) meinte hierzu: „(…) inhaltlich dachte ich mir, Sight-Terp macht das schon (…) wie ein Backup“ und verließ sich somit einerseits auf das Tool, fügte in Hinblick auf den mündlichen Ausdruck und die Präsentationsweise dann

aber hinzu: „(…) Es ist schwierig, den Blickkontakt aufrecht zu halten, und (…) es ist gesprochene Sprache und ich finde so muss es vorgetragen werden (…)“ (TN 10, Pos. 13).

Abschließend wussten die Teilnehmenden nicht konkret, wie technische Geräte, sei es ein Tablet, Handy, oder Laptop, welche Sight-Terp abrufen können, in einem Dolmetscheinsatz, welcher unterschiedliche Formen annehmen kann, im Sitzen, im Stehen, im Gehen eingesetzt werden könnten.

Zusammenfassend kannten die Teilnehmenden die Anwendung davor nicht und würden sich zu Übungszwecken sowie neben dem Erlernen der Notizentechnik auch gerne im Studium eingehender mit Sight-Terp bzw. ähnlichen technologischen Anwendungen auseinandersetzen (vgl. insbesondere TN1, TN2, TN5, TN9, TN10).

Kategorie 3: Konkrete Vorgehensweise mit bzw. ohne Sight-Terp

Die Antworten der Teilnehmenden zeugen insgesamt von einer individuellen Vorgehensweise bei der Verdolmetschung mit bzw. ohne Sight-Terp. Einerseits konzentrierten sich manche TN beim Durchgang mit Sight-Terp eher auf die englische Ausgangsrede und führten eine Vom-Blatt-Übersetzung durch (vgl. TN2, TN4, TN8). Die Mehrheit verwendete andererseits die maschinelle Übersetzung, um die zusätzliche Belastung des Vom-Blatt-Übersetzens zu vermeiden (vgl. TN1, TN3, TN5, TN6, TN7, TN9, TN10). Es entstand dabei jedoch auch die Notwendigkeit umzuformulieren. Beispielsweise sagten TN1 (Pos. 21): „(…) bei geschlechtergerechter Sprache habe ich etwas ändern müssen“, bzw. TN7 (Pos. 13): „(…) natürlich muss man aufpassen, was man da liest (…)“.

Die Notizen wurden mit der Anwendung Sight-Terp von allen Teilnehmenden grundsätzlich in einem kombinierten Modus eingesetzt, wenn auch auf unterschiedliche Art und Weise (vgl. Abschnitt 5.1). Oftmals wurden Problemauslöser, allen voran Zahlen, als die wichtigen Elemente genannt, die in der Anwendung Sight-Terp gesucht wurden. Im Vergleich dazu wurden sprachstrukturelle Elemente, wie Konnektoren, oder auch Adjektiven und Modalverben, meist selbst notiert.

Zusätzlich gaben die Teilnehmenden an, sie hätten zwar weniger notiert (u. a. TN1, TN3, TN9, TN10), jedoch blieben die Notizen oft weiterhin die Grundlage der Verdolmetschung, mit denen dann im Tool auf Vollständigkeit abgeglichen wurde, denn „Ich habe natürlich gemerkt, wo Lücken sind (…) und dann habe ich (…) auf Sight-Terp (…) diese Teile vom Blatt gedolmetscht“ (TN2, Pos. 17; vgl. auch TN4, TN10). Beim Suchen nach der benötigten Information befürchteten manche Teilnehmenden jedoch „ewig lange Pausen“ (vgl. TN4, Pos. 13) bzw. dass sie dabei „ins Stottern“ kommen würden (vgl. TN8, Pos. 14).

Eine ähnliche Herausforderung bei der beschriebenen Vorgehensweise beschrieb TN3 (Pos. 9), denn „(…) dann suche ich vergeblich die Zahl“. TN4 (Pos. 9) meinte ebenso: „(…) einzelne Wörter hätte ich niemals wiedergefunden“.

Auch kognitive Aspekte wurden angesprochen. TN3 (Pos. 17) konzentrierte sich bei der Verdolmetschung mit CAI-Tool sehr auf dieses, „was eben auch die Qualität der Verdolmetschung beeinflussen hätte können oder hat, weil ich mich eben weniger auf die Ausgangsrede konzentriert habe (…) beim Zuhören und Notieren (…)“.

Das Aufrufen von Sight-Terp auf einem mittig aufgestellten I-Pad (vgl. Abschnitt 4.2.2) war außerdem eine Hilfestellung. TN5 (Pos. 13) beschrieb: „(…) das I-Pad habe ich festgehalten (…) und mit meinem Finger bin ich die Notizen durchgegangen (…) das hat schon gut funktioniert“, fragt sich jedoch, wie es wäre „wenn man weniger notiert, ob man sich da nicht verliert“.

Neben dieser Vorgehensweise kontrollierten andere Teilnehmenden die elektronischen Aufzeichnungen von Sight-Terp während des Transkribierens der Ausgangsrede und haben im Falle problematischer Stellen (ausschließlich) diese mitnotiert, um sie in der Verdolmetschung bzw. unter Verwendung der maschinellen Übersetzung korrekt in die Verdolmetschung einzubauen (vgl. TN6, TN9). Hingegen meinte TN8 (Pos. 10), man wäre verleitet, zu nah am Ausgangstext zu bleiben, wenn die englische Transkription als Basis für die Verdolmetschung herangezogen werden würde.

In Bezug auf die Verdolmetschung im Durchgang ohne Sight-Terp behaupteten einige Teilnehmende, dass Symbole und Schlagwörter in den eigenen Notizen eine große Hilfe waren, dahingegen konnte TN7 (Pos. 21) die eigenen Notizen oft nicht mehr entziffern und hatte somit einen schlechteren Eindruck von der eigenen Verdolmetschung ohne Sight-Terp.

Abschließend antwortete TN10 (Pos. 37) auf die Frage, ob mit den reduzierten Notizen aus dem Durchgang nur mit Sight-Terp auch beim plötzlichen Ausfall der Anwendung eine Verdolmetschung möglich gewesen wäre: „Es wäre vielleicht weniger detailliert gewesen, aber es hätte schon funktioniert, denn mein Gedächtnis ist ja auch halbwegs da“ und erinnerte somit daran, dass die Aufnahme technologischer Anwendungen in den Arbeitsprozess des Konsekutivdolmetschens in Hinblick auf ihre Zuverlässigkeit einzuschätzen ist (vgl. Abschnitt 1.2.2).

Kategorie 4: Selbsteinschätzung der Leistung mit bzw. ohne Sight-Terp

Zur Bewertung der eigenen Leistung und ggf. der Qualität ihrer Verdolmetschungen in den beiden Durchgängen mit und ohne Sight-Terp meinten die Teilnehmenden größtenteils, sie

hätten sich mit Sight-Terp verbessert, und zwar aufgrund unterschiedlicher Kriterien, wie Genauigkeit (vgl. TN6, TN9) und auch Flüssigkeit (vgl. TN1). TN5 (Pos. 21) war in der Darbietung nach eigener Aussage viel entspannter, da Sight-Terp im Notfall zur Verfügung gestanden wäre.

Andere Teilnehmende (vgl. TN3, TN9) befürchteten wiederum, dass Sight-Terp Pausen in der Verdolmetschung verursachte bzw. zum Ablesen verleitete und stellten selbst weniger Flüssigkeit im Vergleich zur ausschließlichen Verwendung der Notizen fest. TN10 (Pos. 29) stufte die Verdolmetschung ohne Sight-Terp ebenso als besser ein, denn „ich hatte (…) mehr Kontrolle darüber, was ich sage".

Letztlich sahen TN4 und TN8 keinen Unterschied in den beiden Durchgängen mit und ohne Anwendung und bewerteten beide Verdolmetschungen als gleich gut.

Kategorie 5: Verbesserungsvorschläge

Hier antworteten die meisten Teilnehmenden, dass das Layout der Anwendung Sight-Terp verbessert werden sollte. Es wurden vor allem die fehlerhafte Absatzgestaltung bzw. die zu überarbeitende Nummerierung angesprochen. Die farbigen Markierungen sollten nach TN2 (Pos. 29) und TN10 (Pos. 41) komplett weggelassen werden. Die Markierungen hätten sich andere Teilnehmende neben dem Vorhandensein im englischen Transkript auch in der maschinellen Übersetzung gewünscht (vgl. TN3, TN5, TN8).

TN6 (Pos. 25) sprach sich für „eine übersichtlichere Strukturierung" aus „mit mehr Absätzen und Sinneinheiten", ebenso wie TN9, wobei kürzere Absätze vorteilhafter wären. In diesem Zusammenhang gab TN7 (Pos. 25) zu: „Es ist schwierig (…) das Tool müsste ja dann herausfiltern können was die wichtigsten Informationen sind und das ist nicht leicht zu programmieren". Hingegen führte TN9 (Pos. 41) die falsche Absatzgestaltung auf die rasche Vortragsgeschwindigkeit zurück.

TN8 (Pos. 26) hätte sich neben einem vollständigen englischen Transkript nur einzelne Aufzählungspunkte, also keine ganzen Sätze, in der deutschen Übersetzung gewünscht. Außerdem sollten vor dem Einsatz von Sight-Terp Metainformationen zum Dolmetscheinsatz angegeben werden können, wie das Setting und Sprachregister, etc. (TN1, Pos. 55).

Zum neu entstandenen Arbeitsmodus meinte TN9 (Pos. 45): „Meine Idealvorstellung ist (…), dass (…) gleichzeitig das Englische und das Deutsche erscheinen, um das gleich direkt korrigieren zu können". Dadurch könnte eine korrekte Basis für die konsekutive Verdolmetschung während der Transkription und MÜ durch Sight-Terp mitgestaltet werden.

Kategorie 6: Weitere Bemerkungen

In dieser Kategorie befinden sich zusätzliche Kommentare, die nicht komplett durch die vorherigen abgedeckt werden konnten bzw. die Gedanken abschließend ergänzten und die gemachten Aussagen abrundeten. Da diese Kategorie sehr individuelle Aussagen enthält, die nicht zusammengefasst werden konnten, werden sie einzeln nach jenen Teilnehmenden aufgelistet, welche noch weitere Bemerkungen hinzuzufügen hatten.

TN1 (Pos. 59) fügte im Schlusssatz hinzu: „Ich arbeite sehr gerne mit Tools und ich finde wir sollten mehr in dieser Richtung entwickeln und (…) wir sollten nicht davor Angst haben, dass die Maschinen uns den Job wegnehmen (…). Ich bin überzeugt davon, dass wir Dolmetscher*innen sehr gut mit Tools arbeiten können".

TN2 (Pos. 32) meinte Sight-Terp „(…) ist (…) etwas, das man ausprobieren sollte".

TN4 (Pos. 33) erwähnte, dass man „in den meisten Fällen eine Rede vorab ausgedruckt bekommt" und dass Sight-Terp „dasselbe Prinzip" sei, dass im Unterschied dazu jedoch Hervorhebungen wichtiger Elemente, wie Zahlen, bereits vorhanden wären.

TN5 (Pos. 33) stellte fest, dass „die Technik eine Unterstützung und kein Ersatz" sei und nicht ausschließlich, sondern als weiteres Hilfsmittel herangezogen werden sollte, wenn auch beim Durchgang ohne Sight-Terp wesentlich mehr Stress da war, denn „ich werde nicht nachschauen können im Nachhinein".

TN6 (Pos. 28) fand die Anwendung „überraschend praktisch" und „wenn man weiß wie man es einsetzt, kann man sich viel Arbeit ersparen, aber es übernimmt das Denken nicht". Die „doppelte Überprüfung" konnte TN6 „zum Vorteil nutzen".

In Hinblick auf die Verlässlichkeit kritisierte TN7 (Pos. 9): „Wenn technische Probleme auftreten, kann man sich nicht so gut darauf verlassen" und könnte sich eine ähnliche Anwendung eher beim Simultandolmetschen vorstellen (TN7, Pos. 17).

TN8 (Pos. 29) fragte sich, ob das Sprachenpaar Englisch-Deutsch besser funktioniert als andere Sprachenpaare, insbesondere was die maschinelle Übersetzung betrifft, und befürwortete Sight-Terp für Fälle „wenn man jemanden hat, der ewig nicht aufhört zu reden und wo die Kapazitäten nicht mehr da sind" sowie als Unterstützung für Zahlen.

TN9 (Pos. 9) hob das Üben im Modus des computergestützten Konsekutivdolmetschens für den Einsatz mit Sight-Terp als sehr wichtig hervor.

TN10 (Pos. 40) fragte sich abschließend, wie genau ein technisches Gerät neben dem Notizblock „räumlich" in einem Dolmetscheinsatz Anwendung finden kann, und fasste somit abschließend die Zweifel einiger Teilnehmenden zusammen.

6. Diskussion und Schlussfolgerungen

Das Ziel der vorliegenden Masterarbeit war es, mögliche Antworten auf die Forschungsfragen zu finden, inwiefern die Anwendung Sight-Terp beim computergestützten Konsekutivdolmetschen im Sprachenpaar Englisch-Deutsch im Vergleich zur klassischen Notizentechnik zu einer Qualitätsverbesserung führen könnte. Die Hypothese vermutete eine Verbesserung der Anzahl der ins Deutsche fehlerfrei gedolmetschten Sinneinheiten und Problemauslöser der englischen Redebeiträge unter Verwendung von Sight-Terp. Außerdem sollten die Arbeitsweise und weitere Überlegungen zu Sight-Terp von den Teilnehmenden erfragt werden.

Zu diesem Zwecke wurden neben den grundlegenden Arbeitsmodi beim Dolmetschen auch kognitive Aspekte definiert, wobei das computergestützte Konsekutivdolmetschen (kurz CACI) den Teilnehmenden unter Verwendung Sight-Terp eine weitere Anstrengung, abgekürzt als HMI (kurz für ‚human-machine interaction'), abverlangte (vgl. Gile, 2020).

Das Kapitel über Zugänge zur Qualität erlaubte eine breitere Perspektive auf die Problematik der Qualitätsbeurteilung von Verdolmetschungen. Im Zuge der vorgestellten Kriterien wurde eine Vielzahl möglicher Analyseparametern erkannt, mitunter die Genauigkeit (vgl. Moser-Mercer, 1996: 44), Vollständigkeit (Kalina, 1994), Äquivalenz (vgl. Setton & Motta, 2007) bzw. Sinnübereinstimmung (vgl. Bühler, 1986). Genauigkeit wurde als Aspekt für die Evaluierung der gewonnenen Daten herausgegriffen (vgl. Pöchhacker, 2016: 140f.). Aus den vorgestellten Methoden zu deren Evaluierung wurden die Segmentierung der Verdolmetschungen nach Problemauslösern und Sinneinheiten sowie die von Barik (1975) inspirierte Fehleranalyse nach Zhao (2021) zur Beantwortung der Forschungsfrage angewandt.

Mithilfe des im empirischen Teil gewonnenen und ausgewerteten Datenmaterials soll im Folgenden ein Versuch zur Beantwortung der dieser Arbeit zugrundeliegenden Forschungsfragen unternommen werden.

Als Antwort auf die Forschungsfrage 1 kann gesagt werden, dass die Ergebnisse der ausgewerteten Verdolmetschungen mit bzw. ohne Sight-Terp tendenziell auf eine deutliche Verbesserung mit der Anwendung Sight-Terp im Vergleich zur alleinigen Verwendung der klassischen Notizentechnik hindeuten. Alle Teilnehmenden bis auf TN4 wiesen beim Dolmetschen mit Sight-Terp eine geringere Fehlerquote bei der Verdolmetschung der Sinneinheiten auf. Die korrekte Verdolmetschung der Problemauslöser zeigte bei allen Teilnehmenden außer bei TN3 und TN4, welche eine leicht erhöhte bzw. gleichbleibende Fehlerquote erzielten, ebenso einen höheren Wert bei dem Durchgang mit Sight-Terp.

Außerdem wurde herausgefunden, dass sich TN1, TN3, TN5, TN7 und TN9 anteilsmäßig stärker bei der Verdolmetschung der Rede A2 mit Sight-Terp verbesserten als die jeweils anderen Teilnehmenden, die sich bei der Verdolmetschung der Rede B1 mit Sight-Terp im Vergleich zur Rede A2 ohne Sight-Terp zumeist nicht so deutlich verbesserten. Die durchschnittliche Verbesserung der Sinneinheiten unter Verwendung von Sight-Terp beträgt für jene Teilnehmenden, die die Rede A2 mit Sight-Terp verdolmetschten, 29,7 % und für jene, die die Rede B1 mit Sight-Terp verdolmetschten nur 13,9 %. Für die Problemauslöser betrugen die Verbesserungsquoten 19,2 % für die Verdolmetschungen der Rede A2 mit Sight-Terp und lediglich 17,8 % für die Verdolmetschungen der Rede B1 mit Sight-Terp. Diese Erkenntnis kann ein Anzeichen dafür sein, dass die Rede A2 auch ohne Sight-Terp aufgrund zahlreicher Faktoren, wie jene des Aufbaus, der Inhalte bzw. der Vortragsweise, im Vergleich zur Rede B1 möglicherweise einfacher gedolmetscht werden konnte. Dies könnte auch eine mögliche Erklärung dafür sein, weshalb die Bewertungsergebnisse von TN4 bei dem Durchgang ohne Sight-Terp besser ausfielen. Jedoch stellt diese Annahme keine endgültige Erklärung dar, da TN2 und TN6 ähnlich hohe Verbesserungsquoten erzielten wie TN1 und TN5, die die Rede A2 mit Sight-Terp verdolmetschten.

In Bezug auf die Forschungsfrage 2 ist eine Verlängerung der Dauer der konsekutiven Verdolmetschungen mit der Anwendung Sight-Terp zu erwähnen. Diese Erkenntnis könnte womöglich auf die multiplen Inputs, nämlich das Transkript der Originalrede, die maschinelle Übersetzung und ggf. die angefertigten Notizen der Teilnehmenden zurückzuführen sein, welche eine zusätzliche Koordination erfordern. Diese Annahme wurde im Leitfaden-Interview durch die Aussagen der Teilnehmenden zum Teil auch bestätigt. Zur eigenen Bewertung der konsekutiven Verdolmetschungen mit bzw. ohne Sight-Terp fielen die Antworten durchwachsen aus. Einerseits vermuteten die Teilnehmenden allgemein eine Verbesserung in Hinblick auf die Wiedergabe von Problemauslösern, da Sight-Terp eine zusätzliche Absicherung darstellte, andererseits befürchteten sie eine weniger flüssige Darbietung, gar eine Ablenkung durch das Suchen der erforderlichen Information auf Sight-Terp.

Zusammenfassend kann zur Analyse der durch die Teilnehmenden angefertigten Notizen gesagt werden, dass keine allgemeine Tendenz zur Veränderung der Notation für eine anschließende Verdolmetschung mit bzw. ohne Sight-Terp erkennbar ist. Von den zehn Teilnehmenden sind besonders TN1 und TN9 hervorzuheben, die die Notizen stark reduzierten bzw. die Notationsweisen im Vergleich zur Verdolmetschung ohne Sight-Terp sichtlich veränderten. TN8 benötigte für die Rede B1 mit Sight-Terp etwas weniger Papier, obwohl diese Rede eine höhere Wortanzahl als die Rede A2 ohne Sight-Terp aufwies. Alle anderen

Teilnehmenden wandten in beiden Durchgängen offenbar dieselbe Notationstechnik an und verließen sich nicht einzig und allein bei der Anfertigung der Notizen auf Sight-Terp. Diese Vorgehensweise könnte ein Indiz dafür sein, dass die Studierenden eine leicht abgeneigte bzw. vorsichtige Haltung gegenüber des CAI-Tools aufwiesen (vgl. Fantinuoli, 2018a: 163) bzw. dass die Studierenden, die ihre Notation stark veränderten, besonders positiv demgegenüber eingestellt waren. Außerdem bestätigte sich durchwegs die Erkenntnis von Abuín Gonzàlez (2012: 66), dass Studierende ihre Notizen hauptsächlich in der Ausgangssprache und kaum in der Zielsprache anfertigen. Außerdem haben die Teilnehmenden keine ganzen Sätze bzw. Textabschnitte notiert, sondern lassen aufgrund von Symbolen, Abkürzungen und Zahlen erkennen, dass sie den Fokus auf den Sinn und (meist) weniger auf das Notieren des Wortlautes legten (vgl. Seleskovitch 1968: 161f.).

Die Ergebnisse der im Rahmen dieser Masterarbeit durchgeführten Untersuchung lassen grundsätzlich eine Annahme der beschriebenen Arbeitshypothese zu, wobei die gewonnenen Ergebnisse auch in Hinblick auf mögliche unberücksichtigte Aspekte zu sehen sind, die im Folgenden angesprochen werden.

Die Kritik betrifft zunächst die Tatsache, dass eine beschränkte Anzahl von 10 Teilnehmenden ausgewählt wurde sowie dass diese ebenso wie die beurteilende Person fortgeschrittene Dolmetschstudierende waren, welche noch wenig professionelle Arbeitserfahrung als Dolmetscher*innen aufwiesen. Dieser Umstand ist bei der Durchsicht der Ergebnisse insofern zu berücksichtigen, als dadurch nur ein Teil der Funktionstüchtigkeit von Sight-Terp und der damit möglicherweise erzielten Qualitätsverbesserung im getesteten Sprachenpaar Englisch-Deutsch abgebildet werden konnte. Das herangezogene Redematerial lässt zudem aufgrund einer unterschiedlichen Verbesserungsquote für die Teilnehmenden, die die Rede A2 mit Sight-Terp verdolmetschten (durchschnittlich 29,7 % für die SE und 19,2 % für die PT), im Vergleich zu jenen, die die Rede B1 mit Sight-Terp verdolmetschten (durchschnittlich 13,9 % für die SE und 17,8 % für die PT), eine nicht ganz vollständige Vergleichbarkeit in Hinblick auf den Schwierigkeitsgrad der Reden vermuten.

Aufgrund einer Vielzahl an Bewertungsschemata (vgl. Pöchhacker, 2016: 143) sowie möglichen Qualitätsaspekten ließ die vorliegende Analyse zudem lediglich einen Ausschnitt auf die stets zu definierende Qualität der durch die Teilnehmenden durchgeführten konsekutiven Verdolmetschungen zu. Die Erkenntnis, dass es keine einheitlichen Analysekriterien für die Bestimmung der Qualität einer Verdolmetschung gibt (vgl. Collados Aís & García Becerra, 2015: 368), führte auch in der vorliegenden Arbeit dazu, dass weitere

Beurteilungskriterien, wie etwa die Flüssigkeit oder auch paraverbale Elemente (vgl. Pöchhacker, 1994a: 236), die die Verdolmetschung wesentlich mitformten, unbeachtet blieben.

Zudem gibt die unter Abschnitt 4.2.3 vorgestellte Analysemethode zu erkennen, dass insbesondere das Erkennen von Sinneinheiten sowie deren Beurteilung nach Auslassungen, Ersetzungen und Hinzufügungen aufgrund verschwimmender Grenzen nicht immer eindeutig war (vgl. Seleskovitch & Lederer, 1989: 247). Der Fokus auf das Suchen nach Fehlern ignorierte zudem möglicherweise beabsichtigte translatorische Entscheidungen, wobei sich etwa eine geänderte Reihenfolge oder eine Hinzufügung sehr wohl positiv auf die pragmatische Wirkung der Verdolmetschung auswirken könnten (vgl. Pöchhacker, 2016: 148).

Ein weiterer hier nicht untersuchter Aspekt ist die Frage, wie und ob die Teilnehmenden von Sight-Terp Gebrauch machten. Dieser Frage wurde zwar im Leitfaden-Interview nachgegangen, bei der die Studierenden nach eigener Aussage erzählten, ob und wie sie Sight-Terp verwendeten. Die tatsächliche Verwendung könnte jedoch nur mithilfe spezieller Messmethoden, wie etwa mittels Eye-Tracking, erfolgen (vgl. Fantinuoli, 2018a: 170).

Letzten Endes fiel aufgrund der Neuheit der Anwendung Sight-Terp die Übungsmöglichkeit, die eine Woche betrug, begrenzt aus. Da die Teilnehmenden die klassische Notizentechnik im Rahmen des universitären Studiums bereits länger erprobten, könnte eine fehlende Vertrautheit mit Sight-Terp einen Einfluss auf die Arbeitsweise bzw. die quantitativen und qualitativen Analyseergebnisse gehabt haben.

Künftige Forschungsvorhaben, die sich mit dem computergestützten Konsekutivdolmetschen bzw. genauer mit der Anwendung Sight-Terp befassen, könnten eine andere Sprachenkombination und Direktionalität, etwa von der Fremdsprache in die Erstsprache, berücksichtigen. Bei der Analyse der gewonnenen Daten könnten zahlreiche weitere Qualitätsaspekte, insbesondere die Form betreffend, in den Fokus gerückt werden. Außerdem sollte neben der produktorientierten auch weiter auf der prozessorientierten Ebene geforscht werden (vgl. Fantinuoli, 2018a: 167), um die tatsächliche Arbeitsweise und die Bedürfnisse von Dolmetscher*innen beim computergestützten Dolmetschen zu erschließen.

Wie zu Beginn in der Einleitung erwähnt wurde formen Technologien bereits heute die Zukunft des Dolmetschberufs. Aus diesem Grund bleibt eine wissenschaftliche Auseinandersetzung mit dem computergestützten Dolmetschen auch weiterhin wichtiger denn je.

Bibliographie

Abuín Gonzàlez, Marta (2012). The language of consecutive interpreters' notes: differences across levels of expertise. *Interpreting 14* (1), 55–72. DOI: 10.1075/intp.14.1.03abu.

Ahmed, Safa'a Ahmed (2022). Technology and artificial intelligence in simultaneous interpreting: a multidisciplinary approach. *Occasional Papers 78*, 325–353.

AIIC (2015). *Guidelines for speakers.* https://aiic.org/document/4370/ (Stand: 3.6.2024).

AIIC (2023). *Equipment requirements for (on-site and remote) conference interpreting.* https://aiic.org/company/roster/companyRosterDetails.html?companyId=13416&companyRosterId=120 (Stand: 11.5.2024).

ALPAC (1966). *Language and machines: computers in translation and linguistics: a report*. Washington: National Academy of Sciences, National Research Council.

Barik, Henri C. (1975). Simultaneous interpretation: qualitative and linguistic data. *Language and Speech 18* (3), 272–297. DOI: 10.1177/002383097501800310.

Besacier, Laurent; Barnard, Etienne; Karpov, Alexey & Schultz, Tanja (2014). Automatic speech recognition for under-resourced languages: a survey. *Speech Communication 56*, 85–100. DOI: 10.1016/j.specom.2013.07.008.

Bond, Simon (2024). *Gunning Fog Index*. http://gunning-fog-index.com/ (Stand: 19.5.2024).

Braun, Sabine (2006). Multimedia communication technologies and their impact on interpreting. In: Carroll, Mary; Gerzymisch-Arbogast, Heidrun & Nauert, Sandra (Hrsg.) *Proceedings of the Marie Curie Euroconferences MuTra: Audiovisual translation scenarios Copenhagen*. 1.-5. Mai 2006. http://www.euroconferences.info/proceedings/2006_Proceedings/2006_proceedings.html.

Bühler, Hildegund (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. *Multilingua 5* (4), 231–235. DOI: 10.1515/mult.1986.5.4.231.

Carl, Michael (2021). Translation, artificial intelligence and cognition. In: Alves, Fabio & Jakobsen, Arnt Lykke (Hrsg.) *The Routledge handbook of translation and cognition*. London: Routledge, 500–516. DOI: 10.4324/9781315178127-33.

Carstensen, Kai-Uwe; Jekat, Susanne & Klabunde, Ralf (2010). Computerlinguistik - Was ist das? In: Carstensen, Kai-Uwe; Ebert, Christian; Ebert, Cornelia; Jekat, Susanne; Klabunde, Ralf & Langer, Hagen (Hrsg.) *Computerlinguistik und Sprachtechnologie - eine Einführung*. 3. Aufl. Heidelberg: Spektrum, 1–26.

Čeňková, Ivana (2010). Sight translation. In: Gambier, Yves & van Doorslaer, Luc (Hrsg.) *Handbook of translation studies*. Amsterdam: John Benjamins, 320–323.

Chen, Sijia (2017). Note-taking in consecutive interpreting: new data from pen recording. *The International Journal of Translation and Interpreting Research 9* (1), 4–23. DOI: 10.12807/ti.109201.2017.a02.

Chen, Sijia & Kruger, Jan-Louis (2022). The effectiveness of computer-assisted interpreting: a preliminary study based on English-Chinese consecutive interpreting. *Translation and Interpreting Studies 18* (3), 399–420. DOI: 10.1075/tis.21036.che.

Cheung, Andrew K. F. & Li, Tianyun (2022). Machine aided interpreting: an experiment of automatic speech recognition in simultaneous interpreting. *Translation Quarterly* 104, 1–20. https://www.researchgate.net/publication/367220343_Machine_Aided_Interpreting_An_Ex-periment_of_Automatic_Speech_Recognition_in_Simultaneous_Interpreting.

Collados Aís, Ángela & García Becerra, Olalla (2015). Quality. In: Mikkelson, Holly & Jourdenais, Renée (Hrsg.) *The Routledge handbook of interpreting*. London: Routledge, 368–383. DOI: 10.4324/9781315745381.

Collados Aís, Ángela; Pradas Macías, Macarena; Stévaux, Elisabeth & García Becerra, Olalla (Hrsg.) (2007). *Evaluación de la calidad en interpretación simultánea: parámetros de incidencia*. Granada: Comares.

Corpas Pastor, Gloria (2021a). Language technology for interpreters: the VIP project. In: Esteves-Ferreira, João; Macan, Juliet Margaret; Mitkov, Ruslan & Stefanov, Olaf-Michael (Hrsg.) *Translating and the computer 42*. Genf: Editions Tradulex, 36–48.

Corpas Pastor, Gloria (2021b). Technology solutions for interpreters: the VIP system. *Hermēneus. Revista de traducción e interpretación 23*, 91–123. DOI: 10.24197/her.23.2021.91-123.

Costa-jussà, Marta R. (2018). From feature to paradigm: deep learning in machine translation. *Journal of Artificial Intelligence Research 61*, 947–974.

Dam, Helle V. (2010). Consecutive interpreting. In: Gambier, Yves & van Doorslaer, Luc (Hrsg.) *Handbook of translation studies*. Amsterdam: John Benjamins, 75–79.

Defrancq, Bart & Fantinuoli, Claudio (2020). Automatic speech recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers. *Target 33* (1), 73–102. DOI: 10.1075/target.19166.def.

Desmet, Bart; Vandierendonck, Mieke & Defrancq, Bart (2018). Simultaneous interpretation of numbers and the impact of technological support. In: Fantinuoli, Claudio (Hrsg.)

Interpreting and technology. Berlin: Language Science Press, 13–28. DOI: 10.5281/ZENODO.1493291.

Dillinger, Mike (1994). Comprehension during interpreting: what do interpreters know that bilinguals don't? In: Lambert, Sylvie & Moser-Mercer, Barbara (Hrsg.) *Bridging the gap: empirical research in simultaneous interpretation*. Amsterdam: Benjamins, 155–190.

Ding, Yan Lydia (2017). Using propositional analysis to assess interpreting quality. *International Journal of Interpreter Education 9* (1 (4)), 17–39.

Fantinuoli, Claudio (2009). InterpretBank: Ein Tool zum Wissensmanagement für Simultandolmetscher. In: Baur, Wolfram; Kalina, Sylvia; Mayer, Felix & Witzel, Jutta (Hrsg.) *Übersetzen in die Zukunft: Herausforderungen der Globalisierung für Dolmetscher und Übersetzer: Tagungsband der Internationalen Fachkonferenz des Bundesverbandes der Dolmetscher und Übersetzer e.V*. Berlin: Bundesverband der Dolmetscher und Übersetzer e.V., 411–417.

Fantinuoli, Claudio (2016). InterpretBank. Redefining computer-assisted interpreting tools. In: *Proceedings of the 38th conference Translating and the Computer*. London: Editions Tradulex, 42–52.

Fantinuoli, Claudio (2017). *S*peech recognition in the interpreter workstation. *Translating and the computer 39*. November 2017. https://www.researchgate.net/publication/321137853_Speech_Recognition_in_the_Interpreter_Workstation?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InNpZ251cCIsInBhZ2UiOiJzZWFyY2giLCJwb3NpdGlvbiI6InBhZ2VIZWFkZXIifX0.

Fantinuoli, Claudio (2018a). Computer-assisted interpreting: challenges and future perspectives. In: Corpas Pastor, Gloria & Durán-Muñoz, Isabel. *Trends in E-Tools and Resources for Translators and Interpreters*. Boston: Brill, 153–174. DOI: 10.1163/9789004351790.

Fantinuoli, Claudio (2018b). Interpreting and technology: the upcoming technological turn. In: Fantinuoli, Claudio (Hrsg.) *Interpreting and technology*. Berlin: Language Science Press, 1–12. DOI: 10.5281/ZENODO.1493291.

Fantinuoli, Claudio; Marchesini, Giulia; Landan, David & Horak, Lukas (2022). KUDO Interpreter Assist: automated real-time support for remote interpretation. DOI: 10.48550/ARXIV.2201.01800.

Fantinuoli, Claudio & Prandi, Bianca (2021). Towards the evaluation of automatic simultaneous speech translation from a communicative perspective. 2. Aufl. arXiv. DOI: 10.48550/ARXIV.2103.08364.

García Becerra, Olalla (2015). Order effect, impression formation and their impact on the evaluation of interpreting quality. In: Zwischenberger, Cornelia & Behr, Martina (Hrsg.) *Interpreting quality: a look around and ahead*. Berlin: Frank & Timme, 123–146.

Garzone, Giuliana (2002). Quality and norms in interpretation. In: Garzone, Giuliana & Viezzi, Maurizio. *Interpreting in the 21st century: challenges and opportunities: selected papers from the 1st Forlì conference on interpreting studies, 9-11 November 2000*. Amsterdam: John Benjamins, 107–119.

Gieshoff, Anne Catherine & Albl-Mikasa, Michaela (2022). Interpreting accuracy revisited: a refined approach to interpreting performance analysis. *Perspectives 32* (2), 210–228. DOI: 10.1080/0907676X.2022.2088296.

Gile, Daniel (1983). Aspects méthodologiques de l'évaluation de la qualité du travail en interprétation simultanée. *Meta 28* (3), 236–243. DOI: 10.7202/002899ar.

Gile, Daniel (1985). Le modèle d'efforts et l'équilibre d'interprétation en interprétation simultanée. *Meta 30* (1)*,* 44–48. DOI: 10.7202/002893ar.

Gile, Daniel (1992). Basic theoretical components in interpreter and translator training. In: Dollerup, Cay & Loddegaard, Anne (Hrsg.) *Teaching translation and interpreting: training, talent, and experience*. Amsterdam: John Benjamins, 185–193.

Gile, Daniel (1995/ 2009). *Basic concepts and models for interpreter and translator training*. Amsterdam: John Benjamins.

Gile, Daniel (2016). Experimental research. In: Angelelli, Claudia V. & Baer, Brian James (Hrsg.) *Researching translation and interpreting*. New York: Routledge, 220–228. DOI: 10.4324/9781315707280.

Gile, Daniel (2020). Forty years of Effort models of interpreting: looking back, looking ahead. *Keynote lecture for the Japan Translation and Interpretation Forum 2020 organized by JACI - Japan Association of Conference Interpreters*. August 2020.

Gillies, Andrew (2017). *Note-taking for consecutive interpreting: a short course*. 2. Aufl. New York: Routledge.

Goldsmith, Joshua (2017). A comparative user evaluation of tablets and tools for consecutive interpreters. In: Esteves-Ferreira, João; Macan, Juliet; Mitkov, Ruslan & Stefanov,

Olaf-Michael (Hrsg.) *Translating and the computer 39*. Genf: Editions Tradulex. 40–50.

Goldsmith, Joshua (2018). Tablet interpreting: consecutive interpreting 2.0. *Translation and Interpreting Studies 13* (3), 342–365. DOI: 10.1075/tis.00020.gol.

Gran, Laura; Carabelli, Angela & Merlini, Raffaela (2002). Computer-assisted interpreter training. In: Garzone, Giuliana & Viezzi, Maurizio. *Interpreting in the 21st century: challenges and opportunities: selected papers from the 1st Forlì conference on interpreting studies, 9-11 November 2000*. Amsterdam: John Benjamins, 277–294.

Guo, Meng; Han, Lili & Anacleto, Marta Teixeira (2023). Computer-assisted interpreting tools: status quo and future trends. *Theory and Practice in Language Studies 13* (1), 89–99. DOI: 10.17507/tpls.1301.11.

Hale, Sandra & Napier, Jemina (2014). *Research methods in interpreting: a practical resource*. London: Bloomsbury.

Hamidi, Miriam & Pöchhacker, Franz (2007). Simultaneous consecutive interpreting: A new technique put to the test. *Meta 52* (2), 276–289. DOI: 10.7202/016070ar.

Helfferich, Cornelia (2014). Leitfaden- und Experteninterviews. In: Baur, Nina & Blasius, Jörg (Hrsg.) *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer Fachmedien Wiesbaden, 559–574. DOI: 10.1007/978-3-531-18939-0_39.

Herbert, Jean (1968). *The interpreter's handbook: how to become a conference interpreter*. 2. Aufl. Genf: Georg.

Kalina, Sylvia (1994). Analyzing interpreters' performance: methods and problems. In: Dollerup, Cay & Lindegaard, Anne (Hrsg.) *Teaching translation and interpreting 2: insights, aims, visions*. Amsterdam: John Benjamins, 225–232.

Kalina, Sylvia (2002). Quality in interpreting and its prerequisites - a framework for a comprehensive view. In: Garzone, Giuliana & Viezzi, Maurizio. *Interpreting in the 21st century: challenges and opportunities: selected papers from the 1st Forlì conference on interpreting studies, 9-11 November 2000*. Amsterdam: John Benjamins, 121–130.

Kalina, Sylvia (2015). Measure for measure – comparing speeches with their interpreted versions. In: Zwischenberger, Cornelia & Behr, Martina (Hrsg.) *Interpreting quality: a look around and ahead*. Berlin: Frank & Timme, 15–34.

Kuckartz, Udo & Rädiker, Stefan (2022). Transkriptionsregeln, Hinweise zur automatischen Transkription. In: *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim/ Basel: Beltz Juventa. 199–203.

Lamberger-Felber, Heike (2001). Text-oriented research into interpreting - examples from a case-study. *Hermes, Journal of Linguistics* 26, 39–64.

Lederer, Marianne (2014). *Translation: the interpretive model*. London/ New York: Routledge. DOI: 10.4324/9781315760315.

Luo, Xuanmin (2018). Artificial intelligence and the crisis of translation. *Asia Pacific Translation and Intercultural Studies 5* (1), 1–2. DOI: 10.1080/23306343.2018.1456440.

Mankauskienė, Dalia (2018a). *Problem triggers in simultaneous interpreting from English into Lithuanian*. Summary of doctoral dissertation. Vilnius University Institute of Lithuanian literature and folklore.

Mankauskienė, Dalia (2018b). Problems and difficulties in simultaneous interpreting from the point of view of skill acquisition. Vilnius University. https://www.researchgate.net/publication/325320019_Problems_and_difficulties_in_simultaneous_interpreting_from_the_point_of_view_of_skill_acquisition.

Matyssek, Heinz (2012). *Handbuch der Notizentechnik für Dolmetscher: ein Weg zur sprachunabhängigen Notation; 2 Teile in einem Band*. 2. Aufl. Tübingen: Groos.

Mayring, Philipp (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. 13. Aufl. Weinheim/ Basel: Beltz Juventa.

McCarthy, John; Minsky, Marvin L.; Rochester, Nathaniel & Shannon, Claude E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine 27* (4), 12-14. DOI: 10.1609/aimag.v27i4.1904

Mellinger, Christopher & Hanson, Thomas (2016). *Quantitative research methods in translation and interpreting studies*. London: Routledge. DOI: 10.4324/9781315647845.

Moser-Mercer, Barbara (1996). Quality in interpreting: some methodological issues. *The Interpreters' Newsletter 7*, 43–55.

O'Brien, Sharon (2021). Translation, human-computer interaction and cognition. In: Alves, Fabio & Jakobsen, Arnt Lykke (Hrsg.) *The Routledge handbook of translation and cognition*. London: Routledge, 376–388. DOI: 10.4324/9781315178127.

Orlando, Marc (2014). A study on the amenability of digital pen technology in a hybrid mode of interpreting: consec-simul with notes. *The International Journal of Translation and Interpreting Research 6* (2), 39–54. DOI: 10.12807/ti.106202.2014.a03.

Ortiz, Luis Eduardo Schild & Cavallo, Patrizia (2018). Computer-assisted interpreting tools (CAI) and options for automation with automatic speech recognition. *Tradterm 32*, 9–31. DOI: 10.11606/issn.2317-9511.v32i0p9-31.

Pisani, Elisabetta & Fantinuoli, Claudio (2021). Measuring the impact of automatic speech recognition on number rendition in simultaneous interpreting. In: Wang, Caiwen & Zheng, Binghan (Hrsg.) *Empirical studies of translation and interpreting: the post-structuralist approach*. New York: Routledge, 181–197. DOI: 10.4324/9781003017400.

Pöchhacker, Franz (1994a). Quality assurance in simultaneous interpreting. In: Dollerup, Cay & Lindegaard, Anne (Hrsg.) *Teaching translation and interpreting 2: insights, aims, visions*. Amsterdam: John Benjamins, 233–242.

Pöchhacker, Franz (1994b). *Simultandolmetschen als komplexes Handeln*. Tübingen: Gunter Narr.

Pöchhacker, Franz (2015). Evolution of interpreting research. In: Mikkelson, Holly & Jourdenais, Renée (Hrsg.) *The Routledge handbook of interpreting*. London: Routledge, 62–76. DOI: 10.4324/9781315745381.

Pöchhacker, Franz (2016). *Introducing interpreting studies*. 2. Aufl. London/ New York: Routledge.

Pöchhacker, Franz & Zwischenberger, Cornelia (2010). Survey on quality and role: conference interpreters' expectations and self-perceptions. AIIC. https://aiic.org/document/9646/.

Pradas Macías, Macarena (2023). A review of the evolution of survey-based research on interpreting quality using two models by Franz Pöchhacker. In: Zwischenberger, Cornelia; Reithofer, Karin & Rennert, Sylvi (Hrsg.) *Introducing new hypertexts on interpreting (studies).* (Vol. 160). Amsterdam: John Benjamins, 68–90. DOI: 10.1075/btl.160.04pra.

Prandi, Bianca (2018). An exploratory study on CAI tools in simultaneous interpreting: theoretical framework and stimulus validation. In: Fantinuoli, Claudio (Hrsg.) *Interpreting and technology*. Berlin: Language Science Press, 29–59. DOI: 10.5281/ZENODO.1493293.

Riccardi, Alessandra (2002). Evaluation in interpretation - macrocriteria and microcriteria. In: Hung, Eva. *Teaching translation and interpreting 4: building bridges*. Amsterdam: John Benjamins, 115–126.

Rodríguez González, Eloy; Saeed, Muhammad Ahmed; Korybski, Tomasz; Davitti, Elena & Braun, Sabine (2023). Assessing the impact of automatic speech recognition on remote simultaneous interpreting performance using the NTR model. *International Workshop on Interpreting Technologies SAY IT AGAIN*. 2.-3. November 2023.

Rodríguez, Susana; Gretter, Roberto; Matassoni, Marco; Falavigna, Daniele; Alonso, Álvaro; Corcho, Oscar & Rico, Mariano (2021). SmarTerp: A CAI system to support simultaneous interpreters in real-time. In: *Proceedings of the translation and interpreting technology online conference. Triton 2021*. INCOMA Ltd. Shoumen, Bulgaria, 102–109. DOI: 10.26615/978-954-452-071-7_012.

Rozan, Jean-François (1984). *La prise de notes en interprétation consécutive*. Université de Genève: Librairie de l'université Georg.

Russo, Mariachiara (2010). Simultaneous interpreting. In: Gambier, Yves & van Doorslaer, Luc (Hrsg.) *Handbook of translation studies*. Amsterdam: John Benjamins, 333–336.

Seeber, Kilian G. (2011). Cognitive load in simultaneous interpreting: existing theories — new models. *Interpreting 13* (2), 176–204. DOI: 10.1075/intp.13.2.02see.

Seeber, Kilian G. (2015). Simultaneous interpreting. In: Mikkelson, Holly & Jourdenais, Renée (Hrsg.) *The Routledge handbook of interpreting*. London: Routledge, 79–91. DOI: 10.4324/9781315745381.

Seleskovitch, Danica (1968). *L'interprète dans les conférences internationales: problèmes de langage et de communication*. Paris: Lettres Modernes.

Seleskovitch, Danica (1978). *Interpreting for international conferences: problems of language and communication*. (Stephanie Dailey & E. Norman McMillan, Trans.). Washington, DC: Pen and Booth.

Seleskovitch, Danica & Lederer, Marianne (1989). *Pédagogie raisonnée de l'interprétation*. Paris: Didier Érudition [u. a.].

Setton, Robin & Motta, Manuela (2007). Syntacrobatics - quality and reformulation in simultaneous-with-text. *Interpreting 9* (2), 199–230.

Shlesinger, Miriam (1997). Quality in simultaneous interpreting. In: Gambier, Yves; Gile, Daniel & Taylor, Christopher. *Conference interpreting: current trends in research: proceedings of the international conference on interpreting - what do we know and how? Turku, August 25-27, 1994*. Amsterdam: John Benjamins, 123–132.

Świeczkowski, Damian & Kułacz, Sławomir (2021). The use of the Gunning Fog Index to evaluate the readability of Polish and English drug leaflets in the context of health literacy challenges in medical linguistics: An exploratory study. *Cardiology Journal 28* (4), 627–631. DOI: 10.5603/CJ.a2020.0142.

Tjong Kim Sang, Erik F. & De Meulder, Fien (2003). Introduction to the CoNLL-2003 shared task: language-independent Named Entity recognition. *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, 142–147.

Toury, Gideon (1995). *Descriptive translation studies - and beyond*. Amsterdam: John Benjamins.

Ünlü, Cihan (2023). *Automatic speech recognition in consecutive interpreter workstation: computer-aided interpreting tool 'Sight-Terp'*. Master's thesis. Hacettepe University.

Vik-Tuovinen, Gun-Viol (2002). Retrospection as a method of studying the process of simultaneous interpreting. In: Garzone, Giuliana & Viezzi, Maurizio. *Interpreting in the 21st century: challenges and opportunities: selected papers from the 1st Forlì conference on interpreting studies, 9-11 November 2000*. Amsterdam: John Benjamins, 63–71.

Wahlster, Wolfgang (2000). Mobile speech-to-speech translation of spontaneous dialogs: an overview of the final Verbmobil system. In: *Verbmobil: foundations of speech-to-speech translation*. Berlin/ Heidelberg: Springer, 3–21. DOI: 10.1007/978-3-662-04230-4.

Wang, Xinyu & Wang, Caiwen (2019). Can computer-assisted interpreting tools assist interpreting? *Transletters. International Journal of Translation and Interpreting 3*, 109–139.

Wu, Zhiwei (2019). Text characteristics, perceived difficulty and task performance in sight translation: An exploratory study of university-level students. *Interpreting 21 (2)*, 196–219. DOI: 10.1075/intp.00027.wu.

Yamada, Masaru (2014). Can college students be post-editors? An investigation into employing language learners in machine translation plus post-editing settings. *Machine Translation 29* (1), 49–67. DOI: 10.1007/s10590-014-9167-7.

Zhao, Liuyin (2021). *Simultaneous interpreting with text: a study in the context of the United Nations*. Dissertation. Universität Wien.

Anhang I: Redematerial

Speech A2: 8 ways Japan is prepared for earthquakes

Ladies and gentlemen,
In my previous speech, I talked to you about the biggest earthquake recorded in Japan's history and how Tokyo can be affected by a similar or bigger earthquake.

As Japan experiences earthquakes regularly, they've become one of the best-prepared nations on earth. The ability to innovate, invest, educate, and learn from past mistakes has made Japan the most earthquake-ready country in the world. Now I would like to talk about eight ways Japan prepares for earthquakes.

Firstly, Japan builds earthquake-resistant buildings. All houses are built to resist some level of tremor. Houses in Japan are built to comply with rigorous earthquake-proof standards that have been set by law. These laws also apply to other structures like schools and office buildings. It's said that around 87 % of the buildings in Tokyo can resist earthquakes.

Second of all, phone updates. Every smartphone in Japan is installed with an earthquake and tsunami emergency alert system. Before the impending disasters, the alarm is triggered for around five to ten seconds. The alert system gives users time to quickly seek protection.

Thirdly, Japan's high-speed trains are equipped with an earthquake sensor that are triggered to freeze every moving train in the country if necessary. In 2011, when a 9.0 magnitude quake hit Japan, there were 27 trains in action.

Fourthly, if an earthquake hits the nation, all of Japan's TV channels immediately switch to official earthquake coverage, ensuring that the population is well informed on how to stay safe.

As a fifth measure, schools in Japan run regular earthquake drills once a month. From a young age, schoolchildren are educated on the best way to seek protection and stay safe if an earthquake hits their area.

Another way Japan helps protect its population against future natural disasters is by learning from past events. In 1995, the city of Kobe was struck by a completely devastating earthquake, which killed 5,000 people and destroyed tens of thousands of homes. After the city was rebuilt, the Kobe Earthquake Memorial Museum was constructed in the city.

The next thing that is important for Japan's preparation is the earthquake survival kits. Every household has to keep a survival kit with a flashlight, a radio, a first aid kit and enough food and water to last a few days.

Lastly, the water discharge tunnel is one of the most impressive feats of Japanese engineering. This large and hidden tunnel collects flood waters caused by natural disasters like cyclones and tsunamis and safely redistributes the water into the Edo River. It took 13 years to build this massive tunnel, and it cost 3 billion dollars, but you cannot put a price on how many lives it promises to save.
Thank you.

Speech B1: Violence against women

Ladies and gentlemen,
The European Agency for Fundamental Rights has just published a study about violence against women. The report is based off of interviews with 42,000 women across the 28 Member States of the European Union. I found the results of this study rather interesting.

First of all, there are some horrendous figures. One in three women in Europe has been the victim of violence in adulthood. That means after the age of 15. One in 20 women has been raped and 75% of women have experienced sexual harassment at work. These are terrible figures.

What I found the most striking about this study was a seeming paradox. In Nordic countries, the risk of violence is greater than in southern countries. This is very bizarre, I think. Because typically we see Scandinavian countries as more egalitarian societies. They're not so male-dominated. Equal opportunities have made great advances and the countries' policies reflect this. For example, they have very good childcare provisions in place so that women can return to work after having children. Or, for example, if we look at the Swedish cabinet, more than half of the cabinet is made up of women and women are very well respected in politics.

But if you look at the figures of this study, in Denmark, 52% of women have experienced violence. In Finland, this rate is 47%, whereas in Spain it's 20%.

If we take a look at all of the figures for sexual harassment, 37% of Danish women have experienced sexual harassment compared to 11% in Romania. I found this result very counterintuitive. I would have thought that violence against women was more prevalent in societies that are very patriarchal and where women's roles and women's opinions are secondary

to those of a man. I thought violence against women is more in societies where women are often in a submissive role or perhaps even oppressed.

So, when I tried to think about the reasons behind this strange result, the first and depressing question that I asked was “Does women's success breed resentment?”. Is this some sort of horrible backlash in societies where women have become emancipated? And do men simply feel emasculated and does the media portrayal of successful women in these societies substantiate this to some extent? Is it because women are portrayed as aggressive and power-hungry?

But the study itself suggests a different explanation for these results in Nordic countries. The study says that the emancipation of women in these countries and equal opportunities actually increase the exposure to risk. The authors mean that women work outside the home more in these societies, and they're more likely to go out and meet people. And all of this simply increases your exposure to the risk of violence. Another unconnected factor is urbanization. Countries that are more urbanized see more violence against women.

So, in conclusion, I would say the depressing reality is that whatever the figures, they are much too high. And the real question, to my mind, is what can be done to stop it.
Thank you.

Anhang II: Bewertungs- und Analyseraster

Sinneinheit	Fragewort	Problemauslöser	Kodierung	Verdolmetschung	Fehlerkategorie
Ladies and gentlemen,	Anrede				
The **European Agency** for Fundamental Rights	wer/was	European Agency for Fundamental Rights	E		
has just published a study	was				
about violence against	worüber				
women.	wogegen				
The report is based off	wer/was				
of interviews	worüber				
with **42,000**	wie viel	42,000	Z		
women	wessen				
across the **28**	wie viel	28	Z		
Member States	wessen				
of the **European Union**.	wovon	European Union	E		
I found	wer				
the results	wen/was				
of this study	wovon				
rather interesting.	wie				
First of all,	wie viel	First of all,	Z		
there are some horrendous figures.	was				
One in three	wie viel	One in three	Z		
women	wessen				
in **Europe**	wo				
has been the victim	was				
of violence	wovon				
in adulthood.	wann				
That means after the age	was				
of 15.	wann genau	15	Z		
One in 20	wie viel	One in 20	Z		
women	wessen				
has been raped	was				
and **75%**	wie viel	75%	Z		
of women	wessen				
have experienced	wer				
sexual harassment	wen/was	sexual harassment	T		
at work.	wo				
These are terrible figures.	was genau				
What I found	wer				
the most striking	wie genau				
about this study	worüber				
was a seeming paradox.	was genau				
In **Nordic countries**,	wo	Nordic	E		
the risk	was				
of violence	wessen				
is greater than	wie genau				
in **southern countries**.	wo	southern	E		
This is very bizarre,	was genau				
I think.	wer				
Because typically	warum				
we see	wer				
Scandinavian countries as	wen/was	Scandinavian	E		
more **egalitarian societies**.	wie	egalitarian	T		
They're not so **male-dominated**.	was	male-dominated	T		
Equal opportunities have made	wer/was				
great advances and	was				
the countries' **policies**	wer/was				
reflect this.	was				
For example, they have very good	wer/was				
childcare provisions in place so	was genau	childcare provisions	T		
that women can	wer				
return to work	was				
after having children.	wann				

Or, for example, if we look	wer		
at the **Swedish**	was	Swedish	E
cabinet,	wessen	cabinet	T
more than **half**	wie viel	more than half	Z
of the cabinet is made up of	wessen	cabinet	T
women	wovon		
and women are very well respected	wer genau		
in politics.	wo		
But if you look	wer		
at the figures	was		
of this study,	wessen		
in **Denmark**,	wo	Denmark	A/E
52%	wie viel	52%	Z
of women	wessen		
have experienced violence.	was genau		
In **Finland**,	wo	Finland	A/E
this rate is **47%**,	wie viel	47%	Z
whereas in Spain	wo	Spain	A/E
it's **20%**.	wie viel	20%	Z
If we take a look at all	wer		
of the figures	was genau		
for sexual harassment,	wessen	sexual harassment	T
37% of	wie viel	37%	Z
Danish women	wovon genau	Danish	A/E
have experienced **sexual harassment**	was	sexual harassment	T
compared to **11%**	wie viel	11%	Z
in **Romania**.	wo	Romania	A/E
I found	wer		
this result	wen/was		
very counterintuitive.	wie genau		
I would have thought that	wer		
violence against	was		
women	wogegen		
was more prevalent in	wie genau		
societies that are very **patriarchal**	wo	patriarchal	T
and where women's roles and	wer		
women's opinions are	was genau		
secondary to those	wie viel		
of a man.	wessen		
I thought	wer		
violence against	was		
women	wogegen		
is more in societies	wie genau		
where women are often	wo		
in a **submissive** role	was	submissive	T
or perhaps even oppressed.	was genau		
So, when I tried to think about	wer		
the reasons behind	was		
this strange result,	was genau		
the first and depressing	wie viel	first	Z
question	was		
that I asked was	wer		
'Does women's success	wer/was		
breed resentment?'.	was	breed resentment	T
Is this some sort of horrible **backlash**	was genau	backlash	T
in societies	wo		
where women have become **emancipated**?	was genau	emancipated	T
And do men simply feel **emasculated**	wer	emasculate	T
and does the media portrayal	wer/was		
of successful women	wessen		

in these societies	wo				
substantiate this to some extent?	wie genau				
Is it because women	wer				
are portrayed	was				
as aggressive and power-hungry?	wie genau				
But the study itself suggests	wer/was				
a different explanation	was genau				
for these results	wofür				
in **Nordic countries**.	wo	Nordic	E		
The study says	wer/was				
that the **emancipation**	was	emancipation	T		
of women	wessen				
in these countries	wo				
and equal opportunities	was genau				
actually increase	wie genau				
the exposure to risk.	wessen				
The authors mean	wer				
that women work	was				
outside the home more	wo				
in these societies,	wo genau				
and they're more likely	was				
to go out	was genau				
and meet people.	was genau				
And all of this simply increases	wer/was				
your **exposure to the risk**	was				
of violence.	wovon				
Another unconnected factor	was genau				
is **urbanization**.	wer/was	urbanization	T		
Countries that are more **urbanized** see	wer/was	urbanized	T		
more violence against	was				
women	wogegen				
So, in conclusion,	was				
I would say	wer				
the depressing reality is	was				
that whatever the figures,	was genau				
they are much too high.	wie genau				
And the real question,	was				
to my mind, is	wer				
what can be done to stop it.	was				
Thank you.	Schluss				
Ladies and gentlemen,	Anrede				
In my previous speech,	wann				
I talked to you about	wer				
the biggest earthquake	was				
recorded in **Japan's** history	was genau	Japan	E		
and how **Tokyo** can be affected	was	Tokyo	E		
by a similar	wovon				
or bigger earthquake.	wovon genau				
As **Japan** experiences	wer/was	Japan	E		
earthquakes regularly,	was				
they've become one of	wie viel	one of	Z		
the best-prepared nations	wovon				
on earth.	wo				
The ability	wer/was				
to innovate,	was	innovate	A		
invest,	was	invest	A		
educate,	was	educate	A		
and learn from past mistakes	was genau	learn	A		

has made **Japan**	wen/was	Japan	E		
the most **earthquake-ready country**	was				
in the world.	wo				
Now I would like	wer				
to talk about	was				
eight ways	wie viel	eight	Z		
Japan prepares for earthquakes.	wer/was	Japan	E		
Firstly,	wie viel	Firstly,	Z		
Japan builds	wer	Japan	E		
earthquake-resistant buildings.	was	earthquake-resistant	T		
All houses are built	wer/was				
to resist some level	wofür				
of **tremor**.	wessen	tremor	T		
Houses	wer/was				
in Japan	wo	Japan	E		
are built to comply with	was				
rigorous **earthquake-proof** standards	was genau	earthquake-proof	T		
that have been set by law.	wovon				
These laws also apply to	wer/was				
other structures like	wofür				
schools	was genau				
and office buildings.	was genau				
It's said that around 87%	wie viel	87%	Z		
of the buildings	wessen				
in **Tokyo**	wo	Tokyo	E		
can resist earthquakes.	was				
Second of all,	wie viel	Second of all,	Z		
phone updates.	wer/was				
Every smartphone	wer/was				
in Japan	wo	Japan	E		
is installed with	was				
an earthquake	was genau				
and **tsunami**	was genau	Tsunami	T		
emergency alert system.	was genau	emergency alert system	T		
Before the **impending** disasters,	wann	impending	T		
the alarm is triggered	wer/was				
for around **five to ten**	wie lange	five to ten	Z		
seconds.	wessen				
The alert system	wer/was				
gives users time	was				
to quickly seek protection.	was genau				
Thirdly,	wie viel	Thirdly,	Z		
Japan's high-speed trains	wer/was	Japan	E		
are equipped with	was				
an earthquake sensor that	womit				
are triggered to **freeze**	was genau	freeze	T		
every moving train	wen/was				
in the country,	wo				
if necessary.	wann genau				
In **2011**,	wann	2011	Z		
when a **9.0**	wie viel	9.0	Z		
magnitude quake hit Japan,	was	Japan	E		
there were **27**	wie viel	27	Z		
trains in action.	wessen				
Fourthly,	wie viel	Fourthly,	Z		
if an earthquake	wer				
hits the nation,	was				
all of **Japan's** TV channels	wer	Japan	E		
immediately switch	was				
to official earthquake **coverage**,	was genau				
ensuring that the population	wieso				
is well informed on	wie				

how to stay safe.	worüber		
As a **fifth** measure,	wie viel	fifth	Z
schools	wer/was		
in Japan	wo	Japan	E
run regular earthquake **drills**	was		
once a month.	wie oft	once a month	Z
From a young age,	wann		
schoolchildren are educated	wer		
on the best way	wie		
to seek protection and stay safe	wovon		
if an earthquake hits their area.	wann		
Another way **Japan** helps	wer/was	Japan	E
protect its population	was		
against future natural disasters is	wogegen		
by learning from past events.	wie genau		
In **1995**,	wann	1995	Z
the city of **Kobe** was struck	wer/was	Kobe	E
by a completely devastating earthquake,	wovon		
which killed 5,000	wie viel	5,000	Z
people and	wessen		
destroyed **tens of thousands**	wie viel	tens of thousands	Z
of homes.	wessen		
After the city is rebuilt,	wann		
the **Kobe Earthquake Memorial Museum**	wer/was	Kobe Earthquake Memorial Museum	E
was constructed in the city.	was		
The next thing	wer/was		
that is important for	wie genau		
for **Japan's** preparation	wofür	Japan	E
is the **earthquake survival kits**.	was genau		
Every household	wer/was		
has to keep	was genau		
a survival kit with	was		
a flashlight,	was	flashlight	A
a radio,	was	radio	A
a first aid kit and	was	first aid kit	A
enough food and water to last a few days.	was	enough food and water	A
Lastly, the **water discharge tunnel**	was	water discharge tunnel	T
is one	wie viel	one of	Z
of the most impressive feats	wessen	feats	T
of **Japanese engineering**.	wessen	Japanese	E
This large and hidden tunnel	wer		
collects flood waters caused	was		
by natural disasters	wovon		
like **cyclones**	wovon genau	cyclones	T
and **tsunamis** and	wovon genau	tsunamis	T
safely redistributes the water	was		
into the **Edo River**.	wohin	Edo	E
It took **13**	wie lange	13	Z
years	wessen		
to build this massive tunnel	wen/was		
and it cost **3 billion**	wie viel	3 billion	Z
dollars,	wessen		
but you cannot	wer		
put a price	was		
on how many lives it promises to save.	wofür		
Thank you.	Schluss		

Anhang III: Kodierleitfaden

Kategorie	Definition	Ankerbeispiele	Kodierregeln
K1: Schwierigkeit der Ausgangsreden (Code-Farbe: rot)	kognitive Belastung aufgrund der Eigenschaften der Reden, der Sinneinheiten bzw. verstärkt aufgrund der Problemauslöser in den jeweiligen Reden mit bzw. ohne Sight-Terp sowie auch in Hinblick auf die Vorbereitung der Themen (vgl. dazu Abschnitte 1.2.1, 1.2.2, 2.2.1, 2.2.2)	„Ja also beim ersten Thema hatte ich deutlich weniger Vorwissen und beim zweiten ging es um Menschenrechte und die letzte Modulprüfung, die wir gehabt haben, war über Menschenrechte und daher war es für mich ein bisschen leichter über Menschenrechte zu sprechen, obwohl beide Reden meiner Meinung nach sehr dicht waren“ (TN1, Pos. 3)	die Antwort muss Aspekte zur Beschreibung bzw. Qualifizierung der beiden Reden sowie ggf. zur jeweiligen Vorbereitung darauf bzw. zum abrufbaren Vorwissen enthalten
K2: allgemeine Arbeitsweise mit Sight-Terp (Code-Farbe: violett)	die grundlegenden Eindrücke der Verwendung des CAI-Tools durch die TN (vgl. Abschnitte 3.1, 3.3)	„Ich finde es einfach verständlich, wenn man sich das kurz mal anschaut und sich einliest, kann man gut damit arbeiten und mich hat das nicht abgelenkt“ (TN2, Pos. 3)	die Antwort muss sich allgemein und nur auf das Arbeiten mit Sight-Terp während des Notierens bzw. der Wiedergabe beziehen sowie ggf. auch auf visuelle Eindrücke bzw. die Nützlichkeit für den künftigen Einsatz

K3: Konkrete Vorgehensweise mit/ ohne Sight-Terp (Code-Farbe: blau)	die tatsächliche Verwendung anhand konkreter Beispiele ggf. im Kontrast zur alleinigen Verwendung der Notizentechnik (vgl. Abschnitte 1.2, 3.1, 3.3)	„Aus irgendeinem Grund habe ich das total ausgeblendet, dass da eine deutsche Übersetzung wäre, ich habe mich wirklich auf das Englische konzentriert und nicht, was mir da jetzt halt auf Deutsch auch vorgeschlagen wurde“ (TN4, Pos. 2)	die Antwort soll sich auf die Arbeitsweise anhand spezifischer Anwendungsbeispiele mit Sight-Terp im Gegensatz zur Verdolmetschung ohne Sight-Terp beziehen und die Vorgehensweise im Nachhinein etwas rekonstruieren
K4: Selbsteinschätzung der Leistung bzw. Qualität mit/ ohne Sight-Terp (Code-Farbe: gelb)	eine ad hoc Bewertung der eigenen Leistung bzw. Qualität definiert nach Problemauslösern, im weiteren Sinn auch die Qualität anhand der Sinneinheiten (vgl. Kapitel 2)	„Ich glaube meine Leistung ohne Sight-Terp war besser, weil mir das Video besser gelegen ist, aber ich glaube hätte ich das erste Video ohne Sight-Terp machen müssen wäre es fürchterlich gewesen“ (TN6, Pos. 2)	die Antwort soll auf die Leistung bzw. Qualität und ggf. auch auf die Problemauslöser und Sinneinheiten eingehen und die eigene Verdolmetschung mit/ ohne Sight-Terp selbst beurteilen
K5: Verbesserungsvorschläge (Code-Farbe: grün)	die Frage zielt darauf ab, inwiefern das Tool künftig anders oder	„Die Markierungen in der Ausgangssprache und in der Zielsprache wären auch sehr hilfreich, weil	die Antwort soll einen Einblick in mögliche Kritik geben bzw. Tipps für die künftige

	besser gestaltet werden kann (vgl. Abschnitt 3.3.2)	ich mich eben nur auf diese Spalte konzentrierte aber wie gesagt also ich fände dann allgemein ein besseres Layout nicht schlecht“ (TN3, Pos. 3)	Weiterentwicklung der Anwendung Sight-Terp liefern
K6: weitere Bemerkungen (Code-Farbe: braun)	zusätzliche Kommentare	„Im Endeffekt ist die Technik eine Unterstützung und kein Ersatz, so sehe ich das irgendwie beim Dolmetschen“ (TN5, Pos. 3)	Zusatzkommentare und spannende Überlegungen, die jedoch nicht explizit erfragt wurden fallen in diese Kategorie

Abstract (Deutsch)

Die vorliegende Masterarbeit beschäftigte sich mit dem computergestützten Dolmetschen, genauer gesagt mit dem computergestützten Konsekutivdolmetschen, und umfasste dazu einen methodischen Teil unter Verwendung des von Ünlü (2023) entwickelten Tools Sight-Terp.

Zu diesem Zwecke wurden zuallererst grundlegende Dolmetschmodi präsentiert sowie mit speziellem Fokus auf das Simultan- und Konsekutivdolmetschen die damit einhergehenden kognitiven Aspekte angesprochen. Im Anschluss daran wurden mögliche Analyseperspektiven zur Qualitätsbeurteilung von Verdolmetschungen vorgestellt, darunter die Problemauslöser, Sinneinheiten und die Fehleranalyse. Das folgende Kapitel über computergestütztes Dolmetschen skizzierte die dabei zum Einsatz kommende künstliche Intelligenz anhand der automatischen Spracherkennung und der neuronalen maschinellen Übersetzung.

Der darauffolgende methodische Teil stellte eine Replikation des Experiments von Ünlü (2023) dar, welches auf dem von Ünlü eigens dafür entwickelten Tool Sight-Terp basierte. Das Verfahren umfasste ein sich wiederholendes Messverfahren, das heißt insgesamt wurden 20 Verdolmetschungen von 10 Teilnehmenden analysiert, die zwei Reden jeweils abwechselnd in einem Durchgang mit Sight-Terp und in einem weiteren Durchgang ohne Sight-Terp vom Englischen ins Deutsche konsekutiv dolmetschten. Die Analyse erfolgte anhand der Beurteilung der Genauigkeit erkannter Sinneinheiten und Problemauslöser. Außerdem erlaubte ein an die Verdolmetschungen direkt anschließendes Leitfaden-Interview mit den Teilnehmenden wertvolle Einblicke in die Arbeitsweise und Meinungen zum computergestützten Konsekutivdolmetschen mit der Anwendung Sight-Terp.

Die Ergebnisse zeigten, dass die Verwendung von Sight-Terp die Genauigkeit der gedolmetschten Sinneinheiten und Problemauslöser in beinahe allen Fällen erheblich steigerte. Die Einstellungen der Teilnehmenden verdeutlichten jedoch, dass technische Anwendungen für das computergestützte Dolmetschen nicht zuletzt im Lichte der Verlässlichkeit und Vertrautheit mit dem Tool sowie im Rahmen der Bewertung der Verdolmetschung im Sinne zahlreicher weiterer relevanter Aspekte zu sehen sind, die hier nicht im Analysefokus standen.

Abstract (Englisch)

This master's thesis dealt with computer-assisted interpreting, in particular computer-assisted consecutive interpreting. It included a methodological section using the tool Sight-Terp developed by Ünlü in 2023.

For this purpose, basic interpreting modes were first presented, and the associated cognitive aspects were addressed with a special focus on simultaneous and consecutive interpreting. Second, the chapter on possible perspectives of analysis included problem triggers, units of meaning and error analysis as ways of assessing the quality of an interpretation. The following chapter on computer-assisted interpreting outlined artificial intelligence used in the process based on automatic speech recognition and neural machine translation.

The subsequent methodological section included a replication of Ünlü's experiment (2023), which was based on the tool Sight-Terp developed by Ünlü specifically for this purpose. The procedure comprised a repeated measures design, which meant that a total of 20 interpretations by 10 participants were analyzed, each of whom interpreted two speeches alternately in the consecutive working mode from English into German once with Sight-Terp and once without Sight-Terp. The analysis was based on an assessment of the accuracy of recognized units of meaning and problem triggers. In addition, a guided interview with the participants immediately following the interpretation provided valuable insights into their working methods and opinions on computer-assisted consecutive interpreting with Sight-Terp.

The results showed that the use of Sight-Terp significantly increased the accuracy of the interpreted units of meaning and problem triggers in almost all cases. However, the attitudes of the participants made it clear that technical applications for computer-assisted interpreting should also be seen in the light of reliability, familiarity with the tool as well as in the evaluation of the interpretation in terms of a number of other relevant aspects which were not further analyzed here.