



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Respeaking bei NHK – Live-Untertitelung im
Japanischen Fernsehen“

verfasst von / submitted by

Johanna Russ, BA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2021/ Vienna 2021

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

UA 070 331 378

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Translation Deutsch Japanisch

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Franz Pöchhacker

Danksagung

Obwohl das Schreiben einer Masterarbeit grundsätzlich keine Gruppenarbeit ist, bedeutet dies nicht, dass man diese ganz ohne Unterstützung bestreitet. Ich möchte mich an dieser Stelle ganz besonders bei allen Interview-TeilnehmerInnen bedanken, die sich die Zeit genommen haben, mir in den Interviews wichtige Informationen für diese Arbeit mitzuteilen. Ohne die Zusammenarbeit mit NHK, der Zennancho und der JFD wäre diese Forschung nicht möglich gewesen.

Auch gilt mein Dank meinem Betreuer Univ.-Prof. Mag. Dr. Franz Pöchlhammer für seine stetige Erreichbarkeit und Unterstützung bei Fragen aller Art. Des Weiteren danke ich den KollegInnen aus den Japanisch-Zweigen, die mir nicht nur bei dieser Arbeit, sondern im ganzen Studium immer beiseite gestanden haben.

Zu guter Letzt möchte ich auch noch die moralische Unterstützung meiner Freunde und Familie erwähnen, ohne die ich es sicher nicht durch diesen Prozess geschafft hätte.

Inhaltsverzeichnis

1	Einleitung.....	7
2	Definitionen und Hintergründe.....	12
2.1	Einführung in die Audiovisuelle Translation	12
2.2	Geschichte der Untertitelung	14
2.3	Untertitelungsarten und Definitionen	15
2.3.1	Untertitelung für schwerhörige und gehörlose Personen (SDH).....	16
2.4	Live-Untertitelung	17
2.5	Respeaking	19
2.6	Geschichte und Anwendung des Respeakings	20
3	Forschungsstand zum Thema Respeaking.....	22
3.1	Ausbildung und Kompetenzen	22
3.1.1	Gemeinsamkeiten zwischen Respeaking und Dolmetschen.....	24
3.1.2	Hintergründe der RespeakerInnen	26
3.2	Qualität und Genauigkeit.....	27
3.2.1	Die Wortfehlermodelle	28
3.2.2	Guidelines für Qualität	30
3.2.3	Geschwindigkeit.....	31
3.3	Wörtlichkeit oder Paraphrase	32
3.3.1	Informationsverlust durch Reduktion?	34
3.4	Die Zielgruppe der Gehörlosen und Schwerhörigen	35
3.4.1	Rezeptionsstudien.....	36
4	Respeaking in Japan	39
4.1	Gehörlose in Japan	39
4.2	Nutzung von Spracherkennung zur Live-Untertitelung bei NHK.....	41
4.2.1	Anfänge der Spracherkennung	42
4.2.2	Respeaking-Methode bei NHK	43
5	Methodik.....	51
5.1	Qualitative Interviews	52
5.2	Inhaltsanalyse	55
5.3	Ablauf der Inhaltsanalyse	56
5.4	Durchführung der Forschung	58
5.5	Die Suche nach InterviewpartnerInnen	58
5.6	Fragebogaufbau	61
5.6.1	Fragebogen für NHK.....	61
5.6.2	Fragenbogen für Zielgruppe	63
6	Ergebnisse.....	65
6.1	Qualitative Interviews mit Untertitlern	65

6.1.1	Allgemeine Informationen zur Untertitelung bei NHK und Faith.....	65
6.1.2	Fähigkeiten und Hintergründe der RespeakerInnen und EditorInnen	69
6.1.3	Qualitäts- und formale Kriterien.....	73
6.1.4	Zukunft der Untertitelung in Japan.....	76
6.1.5	Schwierigkeiten und Lösungsansätze.....	77
6.1.6	Fazit und Vergleich	80
6.2	Zielgruppenbefragung	81
6.2.1	Demographische Daten	82
6.2.2	Qualität der Untertitelung.....	83
6.2.3	Formale Kriterien	86
6.2.4	Wörtlichkeit der Untertitel	88
6.2.5	Prozess und Zukunft der Live-Untertitelung.....	89
7	Schlussfolgerungen und Ausblick	92
7.1	Mögliche weitere Forschungsansätze.....	94
Literatur		96
Anhang		99
Anhang 1 Fragebögen		99
Fragebogen Untertitelung Japanisch		99
Fragebogen Untertitelung Deutsch.....		100
Fragebogen Zielgruppe Japanisch		102
Fragebogen Zielgruppe Deutsch.....		103
Anhang 2 Interviewtranskripte.....		104
Interview NHK		104
Interview Faith 1		123
Interview Faith 2		136
Anhang 3 Zielgruppeninterviews		150
Interview T1		150
Interview T2		152
Interview T3		155
Interviews T4 bis T8.....		158
Kurzfassungen.....		173
Kurzfassung Deutsch.....		173
Kurzfassung Englisch.....		174

Abbildungsverzeichnis

Abb. 1: Das WER-Modell (Romero-Fresco & Martinez Perez 2015:30)	28
Abb. 2 WER-Modell II (Romero-Fresco & Martinez Perez 2015:30).....	28
Abb. 3 NER-Model (Romero-Fresco 2015:32).....	29
Abb. 4: Untertitelungssystem (Original Homma 2010:26, Eigene Übersetzung)	42

1 Einleitung

Die Themengebiete der intralingualen Live-Untertitelung und der Nutzung der Respeaking-Methode sind ein eher neues Forschungsgebiet der Translationswissenschaft und auch keine Themengebiete, welche man dieser als Außenstehender vielleicht auf den ersten Blick zuordnen würde. Schließlich findet insbesondere in der intralingualen Untertitelung kein konventioneller Übersetzungsprozess statt. Mit der Frage, inwiefern die Untertitelung dennoch als Teil der Translationswissenschaft gesehen wird, wird sich diese Arbeit noch beschäftigen. Klar ist in jedem Fall, dass die Forschung zu dem mittlerweile als Audiovisuelle Translation bekannten Gebiet seit den 90er-Jahren stetig, und in der letzten Zeit geradezu sprunghaft, zunimmt (Diaz Cintas & Remael 2014, Diaz Cintas et al. 2010, Chaume 2013, Pérez-González 2018, Romero-Fresco 2011).

Innerhalb des Forschungsgebietes der Audiovisuellen Translation beschäftigen sich auch mehr und mehr Arbeiten mit der Verwendung des Respeakings, einer Methode zur Live-Untertitelung, die in den letzten Jahren beim Fernsehen in verschiedenen Ländern vorherrschend geworden ist (Luyckx et al. 2013:3). Die Respeaking-Methode führt zu einem weiteren Zusammenhang mit der Translationswissenschaft und vor allem der Berufsgruppe der DolmetscherInnen. Da die Kompetenzen von (insbesondere Simultan-) DolmetscherInnen Ähnlichkeiten mit jenen aufweisen, die auch für die Arbeit als RespeakerIn notwendig sind, gibt es Bestrebungen, Personen mit Dolmetschausbildung in diesen Bereich einzuführen (vgl. Arumi Ribas & Romero Fresco 2008). Tatsächlich bieten mittlerweile einige Universitäten, darunter auch die Universität Wien, Kurse zur Live-Untertitelung für StudentInnen der Translationswissenschaft an. Die Teilnahme an einem solchen Kurs war auch mein erster Berührungspunkt mit der Respeaking-Methode. In diesem Kurs wurde lediglich intralingual in deutscher Sprache sowie interlingual aus dem Englischen ins Deutsche untertitelt. Mir stellte sich aber von Anfang an die Frage, wie diese Methode in meiner zweiten Arbeitssprache, Japanisch, umsetzbar ist und ob sie tatsächlich eingesetzt wird. Damit begann im Grunde bereits meine erste Recherche für diese Arbeit.

Wie sich herausstellte, wird in Japan neben anderen Methoden durchaus auch Respeaking zur Live-Untertitelung im Fernsehen verwendet. Viele der Herausforderungen, vor denen auch UntertitlerInnen stehen, die in europäischen Sprachen arbeiten, treffen wohl in jedem sprachlichen Kontext zu. Schließlich wird in jedem Fall die Arbeit durch den Live-Kontext und die damit entstehenden Geschwindigkeitsanforderungen geprägt. Dennoch gibt es grundsätzliche Unterschiede zwischen Japanisch und den meisten europäischen Sprachen, die

ein Alphabet nutzen, welches sich stark von dem japanischen Schriftsystem unterscheidet. Von Anfang an faszinierte mich daher die Frage, inwieweit es überhaupt möglich ist, im Japanischen mit Spracherkennung zu arbeiten, und welche besonderen Herausforderungen diese Arbeit mit sich bringt. Eine kurze Recherche im Internet ergab, dass die Respeaking-Methode in Japan bei dem großen staatlichen Sender NHK tatsächlich angewendet wird. Es gibt auch einige in japanischer Sprache abgefasste Artikel dazu, die sich vor allem mit technischen Aspekten beschäftigen. Jedoch war von Anfang an klar, dass ich detaillierte Antworten zur Live-Untertitelung bei NHK nur durch direkten Kontakt mit den zuständigen Personen gewinnen konnte. Da eine Reise nach Japan sich wegen der Covid19-Pandemie leider als unmöglich erwies, wurden als Ersatz für Gespräche vor Ort Telefoninterviews und schriftliche Befragungen als Grundlage für die Arbeit zu verwendet. Nachdem auch tatsächlich einige Kontakte nach Japan zustande kamen, konnte die Forschung auf diese Weise durchgeführt werden.

Im Folgenden wird der Aufbau der Arbeit näher beschrieben, und die Inhalte der einzelnen Kapitel werden erläutert. Grundsätzlich gliedert sich die Arbeit in einen theoretischen und einen empirischen Teil. Der theoretische Teil beinhaltet die Ursprünge der Live-Untertitelung und speziell der Respeaking-Methode sowie den aktuellen Forschungsstand. Dies bezieht sich sowohl auf Arbeiten zum Respeaking allgemein in verschiedenen Ländern, als auch auf jene, die sich speziell mit der Untertitelung bei NHK im japanischen Sprachraum befassen. Im empirischen Teil werden zunächst die verwendeten Methoden und Fragebögen beschrieben, bevor dann die Befragungen ausgewertet werden. Durch diese Aufteilung sollen sowohl die theoretischen Aspekte des Respeakings als auch die praktische Anwendung beim Sender NHK und deren Rezeption durch Zielgruppen adäquat beschrieben werden. Es folgt nun noch eine genaue Inhaltsangabe der einzelnen Kapitel.

Da die Live-Untertitelung in den Themenbereich der Audiovisuellen Translation fällt, wird zunächst dieses Forschungsgebiet näher beschrieben. Hierbei befasst sich die Arbeit mit der Geschichte dieses Bereichs, der relativ neu in der Forschung, nicht jedoch in der Anwendung ist. Speziell werden die Entstehungsgeschichte und die Forschung zur Untertitelung als solche besprochen. Dies dient vor allem dazu, die Live-Untertitelung in einen breiteren Kontext zu setzen. Nachdem die Untertitelung allgemein erläutert worden ist, beschäftigt sich ein weiteres Unterkapitel mit der Live-Untertitelung. Hier soll die Geschichte ihrer Anwendung und Etablierung in den verschiedenen Ländern kurz skizziert werden, außerdem werden die Methoden zur Live-Untertitelung genannt und die bei dieser Art der Untertitelung auftretenden Probleme.

Da Respeaking die Methode ist, auf der in dieser Arbeit der Fokus liegt, wird diese noch einmal ausführlich beschrieben. Hierbei beginne ich mit einer Definition des Respeakings, um dann auf die Anfänge und Geschichte der Methode einzugehen. Anschließend werden auch noch einige der besonderen Charakteristika und Herausforderungen beschrieben. Dies bildet eine gute Überleitung in den Forschungsstand zum Thema Respeaking. Obwohl das Thema der Untertitelung über einen langen Zeitraum hinweg relativ unbeachtet blieb, gibt es nun einen recht breiten Umfang an Forschung dazu. Daher ist es natürlich nicht möglich, in diesem Rahmen alle Arbeiten zu diesem Gebiet zu erwähnen. Allerdings lassen sich die meisten Arbeiten in übergeordnete Themengebiete einteilen, die in einzelnen Unterkapitel abgehandelt werden.

Ein häufiges Thema, das zuvor schon kurz angeschnitten wurde, sind die Kompetenzen und die Ausbildung der RespeakerInnen. Dadurch, dass die Untertitelungsraten auch in der Live-Untertitelung weiter steigen, werden in Zukunft mehr MitarbeiterInnen für diese Aufgabe benötigt werden. Daher ist es von großer Wichtigkeit, herauszufinden, welche Kompetenzen und Ausbildungen diese brauchen, um eine qualitativ hochwertige Untertitelung gewährleisten zu können. Es gibt hierbei mehrere Methoden, die Kompetenzen von RespeakerInnen zu beleuchten. Auf der einen Seite kann man die Expertise der bereits in diesem Bereich arbeitenden Personen als Informationsquelle nehmen (Romero-Fresco 2011, Romero-Fresco 2012). Auf der anderen Seite können Forschende selbst die Herausforderungen der Respeakings untersuchen, oder Experimente dazu durchführen (Arumi Ribas & Romero-Fresco 2008, Geens 2008). Beide dieser Herangehensweisen werden in dem betreffenden Unterkapitel beleuchtet. Außerdem werden bereits existierende universitäre Kurse zu diesem Fachgebiet aufgelistet.

Ein weiterer Bereich, zu dem im Zusammenhang mit Respeaking geforscht wird, ist die Qualität und Genauigkeit der Untertitelung. Dies ist ein sehr wichtiger Forschungsbereich, denn von der Qualität hängt stark ab, wie nützlich Live-Untertitelung wirklich für die Zielgruppen ist. Diese Qualität kann über verschiedene Kriterien gemessen werden. Eine relativ einfache und direkte Methode ist die Ermittlung von Wortfehlerraten (Romero-Fresco 2011). Neben der Anzahl an Fehlern spielt aber auch noch die Geschwindigkeit der Untertitel eine wichtige Rolle. Hierbei geht es besonders um die Verzögerung zwischen dem, was auf dem Fernsehbildschirm gezeigt wird, und der Umsetzung in Untertitel (Romero-Fresco 2009). Ein weiterer Aspekt, der mit Qualität zu tun hat, sind die Qualitätsrichtlinien, auf die sich die einzelnen Fernsehsender beziehen und die ebenfalls in diesem Kapitel behandelt werden.

Neben jenen zur Qualität als solcher sind auch fundamentale Fragen danach zu beantworten, was eine gute Untertitelung überhaupt ausmacht. Eine Fragestellung, die besonders häufig in diesem Zusammenhang diskutiert wird, ist die Wörtlichkeit der Untertitel. Hier müssen UntertitlerInnen einen oftmals schwierigen Spagat zwischen einer möglichst wortgetreuen Wiedergabe des Gesprochenen und der Qualität und Geschwindigkeit der Untertitel bewältigen. Der Grund, weshalb eine wörtliche Wiedergabe auch in schwierigen Situationen oft bevorzugt wird, ist die Gefahr eines Informationsnachteils jener Personen, die auf die Untertitelung angewiesen sind. Daher muss man sich auch damit befassen, ob und inwiefern ein solcher stattfindet (Luyckx 2013, Romero Fresco 2009).

Zudem sind UntertitlerInnen mit der Schwierigkeit konfrontiert, dass die Zielgruppe, für die sie arbeiten, sehr divers ist. Nicht nur richtet sich Untertitelung nicht ausschließlich an Gehörlose und Schwerhörige, obwohl diese auch in dieser Arbeit als die Hauptzielgruppe betrachtet wird, auch innerhalb dieser Gruppen sind die Erwartungen und Erfahrungen sehr unterschiedlich (Romero Fresco & Martinez Perez 2015). Auch mit dem Thema der Zielgruppen und deren Rezeption der Untertitel befasst sich diese Arbeit. Diese Themenbereiche bilden eine wichtige Grundlage insbesondere auch für die empirische Forschung.

Ehe diese beschrieben wird, soll zunächst allgemein die Situation in Japan erläutert werden. Hierzu werden zunächst einige Fakten zur Untertitelung in Japan erläutert und auf die Frage eingegangen, wie die Situation der Gehörlosen und schwerhörigen Personen in Japan aussieht. Abgesehen von jenen Personen, die in jungen Jahren schon (teilweise) ihr Gehör verloren haben, spielt hier auch Japans Status als überalterte Gesellschaft eine Rolle. Außerdem wird die gesetzliche Lage erklärt und den Fragen nachgegangen, welche Vorgaben die japanische Regierung für die Untertitelung bei Live-Sendungen gibt, und wie die Fernsehsender versuchen, diese umzusetzen. Diese kleine Einführung bietet einen allgemeinen Hintergrund zur Situation und macht die Darstellung der Untertitelungssituation in den nachfolgenden Kapiteln deutlicher.

Im Anschluss an diese Einführung werden die Geschichte und Anwendung der Untertitelung bei NHK, insbesondere in technischer Hinsicht, aus den verfügbaren Arbeiten heraus beschrieben. Hierbei wird der Fokus zwar wieder auf das Respeaking gelegt, die Arbeiten beschreiben aber auch die Anwendung von anderen Methoden. Es werden hier auch schon erste Schwierigkeiten und Lösungsansätze in der Live-Untertitelung beschrieben, die später als Ansatz für die tiefer gehenden Interviewfragen dienen. Thema ist auch die Auswirkung der Erdbeben-Katastrophe des Jahres 2011, die in einer tragischen Weise als

großer Impetus für den Fortschritt der Live-Untertitelung gewirkt hat (Homma 2010; Satou 2018; Ito 2013). Nach diesem letzten Theorieabschnitt folgt der empirische Teil der Arbeit.

Für den empirischen Teil werden als Erstes die Methodik der Arbeit vorgestellt. Es handelt sich bei meiner Vorgehensweise um eine qualitative Forschung und die Gründe für diese Wahl werden zunächst dargelegt. Genauer betrachtet kommt hier das qualitative Interview zum Einsatz, allerdings mit einigen Sonderfaktoren, da die Interviews aufgrund der räumlichen Entfernung der InterviewpartnerInnen und der Unmöglichkeit von Reisen komplett elektronisch durchgeführt wurden. Im Falle der Zielgruppenbefragung, bei der der weitere Faktor einer Sprachbarriere dazukam, wurden diese Interviews sogar nur schriftlich geführt. Ausgewertet wurden die Interviews anschließend durch eine qualitative Inhaltsanalyse nach Mayring. Auch diese Analysemethode wird im Methodikkapitel vorgestellt. Ebenso wird der Ablauf der Interviews beziehungsweise die Suche nach InterviewpartnerInnen kurz beschrieben. Hierzu wurden Interviewleitfäden in Form von Fragebögen erstellt, welche ebenfalls inhaltlich präsentiert werden. Die Grundlage für die an die UntertitlerInnen gerichteten Fragen bildet die gesamte zuvor recherchierte Literatur zum Thema Respeaking, insbesondere aber auch zur Live-Untertitelung in Japan. Bei den Fragen, die sich an die Zielgruppe der Gehörlosen und Schwerhörigen richten, dienten hauptsächlich Rezeptionsstudien als Grundlage.

Das anschließende Kapitel beschäftigt sich mit den Ergebnissen der Auswertung der Interviews. Hierbei wird zunächst die Befragung der UntertitlerInnen ausgewertet. Wie dies auch schon im Fragebogen der Fall war, orientiert sich die Aufteilung der Auswertung an den großen Themengebieten des Respeakings und ist in die Unterkapitel „Allgemeine Informationen“, „Fähigkeiten der UntertitlerInnen“, „Qualitäts- und formale Kriterien“, „Schwierigkeiten und Lösungsansätze“ sowie „Zukunft der Live-Untertitelung in Japan“ gegliedert. Darauf folgt die Auswertung der Zielgruppeninterviews, welche sich ebenfalls in thematisch gegliederte Unterkapitel aufteilt. Besprochen werden die demographischen Daten der interviewten Personen und deren Fernsehgewohnheiten, Meinungen zu Qualität und den formalen Kriterien der Untertitelung, die Frage nach der Wörtlichkeit der Untertitel sowie der Prozess und die Zukunft der Live-Untertitel.

Es folgt eine Zusammenfassung der wichtigsten Punkte in einem Fazit. Abgerundet wird die Arbeit durch eine Conclusio, in der die Ergebnisse noch einmal zusammengefasst und in Beziehung zur Literatur gesetzt werden. Es folgt auch noch eine Beleuchtung der Limitationen dieser Studie, die nicht außer Acht gelassen werden dürfen. Darüber hinaus wird aufgezeigt, welche weiterführenden Forschungen zu dieser Thematik noch möglich wären.

2 Definitionen und Hintergründe

In diesem Unterkapitel werden einige grundlegende Themen der Untertitelung und audiovisuellen Translation im Allgemeinen besprochen, um eine Einführung in die Forschungsbereiche zu geben und das Thema Live-Untertitelung und Respeaking mit seinem Hintergrund betrachten zu können. Auch sollen einige für die Arbeit wichtige Begriffe definiert werden. Es folgt daher eine kurze Beschreibung jenes Forschungsgebiets, in dessen Kontext sich diese Arbeit einbetten lässt, der Audiovisuellen Translation (AVT). Hierbei handelt es sich selbstverständlich nur um einen kurzen Abriss dieses sehr umfangreichen Themas, um den Rahmen der Arbeit nicht zu sprengen. Außerdem wird die Untertitelung in ihren verschiedenen Varianten und mit ihrem geschichtlichen Hintergrund besprochen. Von der allgemeinen Thematik ins Spezifische übergehend, befasst sich das Unterkapitel schließlich noch mit der Live-Untertitelung und dem Respeaking, wozu Definitionen und allgemeine Informationen ausgeführt werden.

2.1 Einführung in die Audiovisuelle Translation

Wie bereits erwähnt, lässt sich das Thema (Live-)Untertitelung im translationswissenschaftlichen Kontext in den Bereich der Audiovisuellen Translation (engl. *audiovisual translation*) einordnen (Perez-Gonzalez 2018). Diesem Forschungsgebiet wurde trotz der langjährigen Präsenz translatorischer Aktivität bis vor kurzer Zeit wenig Aufmerksamkeit geschenkt. Seit Mitte der 90er-Jahre hat die Forschung zu audiovisuellen Themen aber einen festen Platz in der Translationswissenschaft erhalten und wird nunmehr als einer der am schnellsten wachsenden Forschungszweige in der Translationswissenschaft bezeichnet (Perez-Gonzalez 2018, Diaz Cintas & Baños Piñero 2016, Diaz Cintas et al. 2010).

Seit ihrer Entstehung hat die Translationswissenschaft einige Paradigmenwechsel oder „turns“ erlebt. Die stärkere Hinwendung zu audiovisuellen Medien als Forschungsobjekt wird daher auch häufig als der „audiovisual turn“ bezeichnet. Es sollte allerdings erwähnt werden, dass nicht alle Forschenden mit diesem sehr vereinfachten Begriff des „turns“ als Bezeichnung für eine komplexe Entwicklung im Wechselspiel von Faktoren wie beispielweise Technologie und Globalisierung übereinstimmen (Perez-Gonzalez 2018:1). Klar ist, dass es eine Entwicklung gegeben hat, die unter anderem mit den rasanten technischen Neuerungen auf dem audiovisuellen Gebiet zusammenhängt, welche eine große Auswirkung auf das tägliche Leben der Bevölkerung mit sich bringen. Hinzu kommt ein größer werdendes Interesse an Barrierefreiheit als treibende Kraft hinter diesen Weiterentwicklungen (Remael et al. 2016:249)

Das Themenfeld der audiovisuellen Translation umfasst, grob zusammengefasst, die Bereiche inter- und intralinguale Untertitelung, Synchronisation, Voice-over und Audiodeskription. Alle diese Methoden können in verschiedenen Kontexten wie Film, Fernsehen oder aber auch Theateraufführungen, Konzerten und anderen Umgebungen eingesetzt werden. Als sich der Forschungszweig zum ersten Mal etablierte, wurde neben dem Namen des Forschungsfeldes, für das es verschiedenste Ausdrücke gab, auch sein Platz in und Verhältnis zur Translationswissenschaft diskutiert. Schließlich ist die Audiovisuelle Translation keine konventionelle Disziplin der Translation. Gottlieb spricht von einer „diagonalen Translation“, bei der ein Text nicht nur in eine andere Sprache übertragen, sondern überdies verschriftlicht wird, im Gegensatz zur horizontalen Translation, in welcher nur der erste Schritt stattfindet. Diese Art der Übersetzung wird von Gottlieb als intersemiotisch bezeichnet (Munday 2016:277). Einen der ersten Schritte, um die audiovisuelle Translation zu integrieren, machte Reiß (1977:99) mit ihrer Definition sowie Integration des audio-medialen Texttyps in ihre Texttypenlehre. Sie definiert diesen Texttyp als eine Art Superstruktur, die über den anderen drei Texttypen steht.

Wegen der durch die verschiedenen Modi auftretenden Längenbeschränkungen ist die Untertitelung somit einer Einschränkung bei der Zeichenanzahl unterworfen (Perez-Gonzales 2018:16). Aufgrund dieser Einschränkungen wurde die Untertitelung auch teilweise als eine eingeschränkte Übersetzung und damit als „Adaption“ bezeichnet, wobei Gambier den Begriff „transadaption“ prägte (Munday 2016:278). Durch diesen Begriff der Adaption wurde diese Praxis als unterschiedlich zur „eigentlichen“ Übersetzung aufgefasst und daher in der Forschung über längere Zeit wenig beachtet (Diaz Cintas & Remael 2014:9). Ein weiterer Diskussionsansatz befasste sich damit, ob audiovisuelle Translation eine eigene Disziplin sein sollte oder sich der Filmübersetzung unterordnen sollte (Diaz Cintas 2009).

Heute ist klar, dass der Translationsbegriff flexibler aufgefasst werden muss, damit er nicht zu einseitig wird und damit eventuell komplett an der Realität vorbei geht (Diaz Cintas & Remael 2014:9f). So ist auch die intralinguale Untertitelung, obgleich bei dieser kein traditioneller Übersetzungsprozess stattfindet, mittlerweile zweifellos Gegenstand verschiedener Forschungen innerhalb der Translationswissenschaft. Trotz der anfangs schwierigen Einordnung hat sich die audiovisuelle Translation nun als fixer Bestandteil der Translationswissenschaft und als eigener wachsender Forschungszweig etabliert. Diese Arbeit wird sich nicht mit allen Bereichen der Audiovisuellen Translation befassen, daher soll im nächsten Kapitel nur eine kurze Geschichte der Untertitelung als Teilgebiet umrissen werden.

2.2 Geschichte der Untertitelung

Die Geschichte der Untertitelung kann paradoxerweise als mit den ersten Stummfilmen beginnend beschrieben werden. Da der Dialog für die Zuseher nicht hörbar war, wurden wichtige Aspekte der Erzählung auf sogenannten Intertiteln zwischen den Szenen eingeblendet. Diese Zwischentitel werden oft als eine Vorgänger-Form der Untertitel betrachtet, obgleich sie einen anderen Zweck als diese erfüllten und neben Dialogen auch narrative Elemente wiedergaben. Die Zwischentitel wurden übersetzt, öfters aber auch editiert und geschnitten. In der Forschung wurde der Ära des Stummfilmes bisher eher wenig Aufmerksamkeit geschenkt, da diese Form der Übersetzung als unproblematisch galt. (O’Sullivan & Cornu 2018:15, Guillot 2018:32).

Mit dem Aufkommen des Tonfilms mussten allerdings andere Methoden zur Internationalisierung des Films gefunden werden. Da am Anfang weder Untertitelung noch Synchronisierung etabliert waren, wurden Filme oft einfach in der Originalsprache, also zumeist Englisch, gezeigt oder sie wurden sozusagen zu Stummfilmen „zurückeditiert“ und wieder mit Intertiteln gezeigt. So wurden bereits zur selben Zeit Methoden entwickelt, die Filme durch separate Projektionen auf die Leinwand zu „untertiteln“ oder aber sogar zu synchronisieren. Einige bedeutende Hollywoodstudios setzten bereits in den 30er-Jahren Synchronisierung ein (O’Sullivan & Cornu 2018:15f).

Bereits zu einem relativ frühen Zeitpunkt kristallisierten sich unterschiedliche Zugänge heraus, Filme zu übersetzen. So gab es Länder, in denen Untertitelung als präferierte Methode überwog, und jene, die Synchronisation bevorzugten. So waren die skandinavischen Länder, die Niederlande, Belgien, Portugal und Griechenland traditionelle Untertitelungsländer, während Länder mit einer großen eigenen Filmindustrie wie Deutschland, Frankreich und Italien eher Synchronisation anwandten. Diese Einteilung hat zum Teil bis heute Bestand (Remael et al. 2016). Die Synchronisation war besonders anfangs mangels guter technischer Möglichkeiten zur Tonabmischung ein schwieriges Unterfangen und wurde vor allem von den großen Filmproduzenten Deutschland und den USA praktiziert. Zu Beginn war aber auch die Untertitelung der Filme technisch betrachtet ein schwieriger Prozess mit oft eher minderwertigen und schwer lesbaren Ergebnissen. Es gibt leider heute kaum Quellen, die die frühe Untertitelung genauer beschreiben und auch wenig Information darüber, wer die UntertitlerInnen waren (O’Sullivan & Cornu 2018:20f).

Mit der Einführung der elektronischen Untertitel im VHS-Format in den 70er-Jahren wurde der technische Prozess der Untertitelung vereinfacht, umfasste allerdings immer noch

mehrere komplexe Arbeitsschritte, um die Untertitel in die richtige Position zu bringen. Heute hingegen sind die Abläufe zwar immer noch ähnlich, durch die Digitalisierung können sie jedoch mithilfe von Untertitelungsprogrammen wesentlich einfacher und für beinahe jeden zugänglich eingefügt werden. Dadurch ist es auch Laien möglich geworden, nach Belieben Untertitelungen anzufertigen (Guillot 2018: 32).

Durch die enge Koppelung zwischen Film und der frühen Entwicklung der audiovisuellen Translationsmethoden in diesen Bereich ist auch der Fachbereich der audiovisuellen Translation eng an die Filmforschung gebunden. Heute befasst sich die AVT mit politischen, aber auch mit technischen Themen, sowohl historisch gesehen als auch mit Gegenwartsbezug, und bedient sich dabei Methoden aus der Filmwissenschaft, quantitativen Methoden aber auch Fallstudien (O’Sullivan/Cornu 2018).

2.3 Untertitelungsarten und Definitionen

In diesem Unterkapitel soll die gegenwärtige Situation der Untertitelung dargestellt werden, beziehungsweise eine Erläuterung versucht werden, was man eigentlich genau unter Untertitelung versteht. Remael und Diaz Cintas definieren den Begriff Untertitelung wie folgt:

Subtitling may be defined as a translation practice that consists of presenting a written text, generally on the lower part of the screen, that endeavours to recount the original dialogue of the speakers, as well as the discursive elements that appear in the image (letters, inserts, graffiti, inscriptions, placards, and the like), and the information that is contained on the soundtrack (songs, voices off).“ (Diaz Cintas & Remael 2014:8)

Diese Definition umfasst ein breites Spektrum an Unterarten und Anwendungsgebieten. Generell lassen sich verschiedene Arten von Untertiteln auf der linguistischen, der zeitlichen, der technischen und der Distributionsebene unterscheiden. Linguistisch betrachtet können Untertitel in die Kategorien intralinguale und interlinguale Untertitelung unterteilt werden, je nachdem ob die Übertragung innerhalb einer Sprache oder von einer Sprache in eine andere Sprache stattfindet. Unterschiedlich sind zwischen diesen beiden Kategorien auch die Hauptzielgruppen, da eine intralinguale Untertitelung normalerweise auf Gehörlose und schwerhörige Personen abzielt. Diese Zielgruppen sind aber natürlich nicht in Stein gemeißelt, da auch interlinguale Untertitelung sich an gehörlose Menschen richten kann, und auch hörende Personen aus verschiedenen Gründen intralinguale Untertitel verwenden (Diaz Cintas & Remael 2014:14f).

Auf der zeitlichen Ebene unterscheidet man nach der Vorbereitungszeit, die die UntertitlerInnen für die Erstellung der Untertitel zur Verfügung haben. An den beiden Enden des Spektrums unterscheidet man zwischen vorbereiteten Untertiteln und Live-Untertiteln. Es gibt allerdings auch Untertitelungssituationen, die einen hybridartigen Charakter aufweisen, oder die zwar vorbereitet sind, aber dennoch einer starken Zeiteinschränkung unterliegen („Semi-Live“) (Diaz Cintas & Remael 2014:19).

Aus technischer Sicht wiederum kann Untertitelung entweder als *open captioning*, bei dem die Untertitel in jedem Fall sichtbar sind, oder als *closed captioning*, in welchem der Zuseher die Möglichkeit hat, die Untertitel zu aktivieren oder zu deaktivieren, konzipiert werden. Traditionell waren die interlingualen Untertitel eher mit *open captioning* und die intralingualen mit *closed captioning* assoziiert, jedoch ist diese Unterscheidung heutzutage nicht mehr in Stein gemeißelt (Munday 2016:278, Diaz Cintas & Remael 2014:22).

Letztlich spielt auch die Distribution eine Rolle. Werden die Untertitel übers Fernsehen ausgestrahlt oder auf eine DVD gebrannt? Neben Filmen und Fernsehserien kann Untertitelung außerdem noch bei kulturellen Ereignissen eingesetzt werden, wie zum Beispiel im Theater oder bei Konzerten (Diaz Cintas & Remael 2014:23). Es gibt abgesehen von dieser sehr groben Unterteilung selbstverständlich noch viele Unterkategorien und kategorienübergreifende Verwendungen, die allerdings den Rahmen dieser Arbeit sprengen würden. Besonderer Fokus der weiteren Kapitel ist die zeitliche Kategorie der Live- oder auch Echtzeit-Untertitel, zu denen im Folgenden noch eine Einführung gegeben wird. Darüber hinaus ist auch die Sonderform der Untertitelung für schwerhörige und gehörlose Personen in dieser Arbeit wichtig.

2.3.1 Untertitelung für schwerhörige und gehörlose Personen (SDH)

Bei der Untertitelung für Schwerhörige und Gehörlose handelt es sich um eine besondere Form der Untertitelung, bei der auch nichtsprachliche auditive Elemente wie Hintergrundgeräusche miteinbezogen werden müssen. Untertitelung speziell für diese Personen wird auch als *Subtitling for the deaf and hard of hearing* (SDH) bezeichnet (Romero-Fresco 2009:110, 2015:10). Wie auch im Bereich des Respeakings ist die BBC im SDH ein Vorreiter und begann bereits in den 1970ern, solche Untertitel im Fernsehen anzubieten. Als nächste Länder zogen in den 1980ern der flämischsprachige Teil von Belgien, Deutschland, Italien und die Niederlande nach. Die Entwicklung von SDH im Fernsehen hängt auch stark mit der Verwendung des Teletexts zusammen, die ebenfalls in vielen Ländern in den 80ern erstmals eingeführt wurde. Am Anfang war nur ein sehr kleiner Anteil der Sendungen untertitelt,

meistens solche, die als besonders von Interesse für die Zielgruppe empfunden wurden. Die Entwicklung von SDH und deren Geschwindigkeit ist eng verbunden mit den Bestrebungen von Gesetzgebern auf der einen und dem Druck durch Gehörlosen-Organisationen und Verbänden auf der anderen Seite. Auch ist zu bemerken, dass die kommerziellen TV-Sender normalerweise hinter den öffentlichen Sendern hinterherhinken, was die Erweiterung von SDH betrifft (Remael 2007:24f).

Während Kampagnen für mehr Zugang zu Medien und Information lange Zeit Aufgabe einzelner humanitärer Organisationen waren, spielen nun wie bereits erwähnt auch legislative Kräfte wie europäische Institutionen eine große Rolle (Díaz Cintas & Remael 2014). Die Weiterentwicklung der Untertitelung in den einzelnen Ländern ist meistens die Folge neuer Legislationen oder anderer Übereinkommen zwischen Regierungen und öffentlichen Rundfunkanstalten (Romero-Fresco 2011:9). Beispielsweise führen die Vereinten Nationen in der *UN-Konvention über die Rechte von Menschen mit Behinderungen* explizit das Recht von behinderten Personen auf „Zugang zu Fernsehprogrammen, Filmen, Theatervorstellungen und anderen kulturellen Aktivitäten in zugänglichen Formaten haben“ an (BMASGK 2016:25). Auch die Europäische Union hat basierend auf der Konvention der Vereinten Nationen die *Europäische Strategie zugunsten von Menschen mit Behinderungen 2010–2020* herausgegeben, die den Zugang zu audiovisuellen Mediendiensten wie Fernsehsendungen und damit verbundenen Verbraucherendgeräten einschließt (European Commission 2021). Diesem Druck ausgesetzt, wird die Live-Untertitelung auch in Zukunft im Namen der Barrierefreiheit aufrechterhalten und auch ausgeweitet werden müssen.

2.4 Live-Untertitelung

Die meisten Menschen denken bei dem Wort „Untertitelung“ vermutlich als Erstes an vorbereitete Untertitel, die beispielsweise für auf DVD gebrannte Filme vorhanden sind und dann ausgewählt werden können. Auch könnte der erste Gedanke eher den interlingualen Untertiteln gelten. Natürlich brauchen aber auch live aufgenommene Sendungen Untertitel, um für alle Personengruppen barrierefrei zugänglich zu sein. Diese Untertitel sind normalerweise intralingual, auch wenn es mittlerweile auch interlinguale Live-Untertitelung gibt (Romero-Fresco & Pöchhacker 2017). Diese intralinguale Live-Untertitelung – im englischen Sprachraum auch als *real-time captioning* bekannt – wird von der International Telecommunication Union bezeichnet als: “A real-time transcription of spoken words, sound effects, relevant musical cues and other relevant audio information in live [...] events” (ITU

2018: 2). Da sie als ein essentielles Werkzeug für eine barrierefreie Teilnahme aller Menschen an der audiovisuellen Kommunikation erachtet werden, sind Untertitel für Gehörlose und Hörgeschädigte in den letzten Jahren in den Fokus von Forschenden, Nutzerverbänden und öffentlichen Stellen gerückt (Romero-Fresco 2015:56).

Bei der Einführung und den Methoden der Live-Untertitelung gibt es große Unterschiede in den einzelnen Ländern. In den USA begann das National Captioning Institute bereits 1982 mit einem Live-Untertitelungs-Service. Für diesen wurden StenotypistInnen, die ursprünglich Gerichtsprozesse mitprotokollierten, umgeschult, um in Zukunft Live-Fernsehprogramme zu untertiteln. In Europa fanden die ersten Experimente in jenen Ländern statt, die auch beim Respeaking als Vorreiter gelten, nämlich in Großbritannien und Belgien. Im Jahr 1982 experimentierte der britische Sender ITV zum ersten Mal mit Live-Untertitelung, wobei damals ein herkömmliches QWERTY-Tastatur genutzt wurde, um große öffentliche Veranstaltungen wie einen Besuch des Papstes oder wichtige Fußballspiele zu untertiteln. Mit einer herkömmlichen Tastatur konnten allerdings nicht die benötigten Geschwindigkeiten erreicht werden, daher begann ITV im Jahr 1987 das sogenannte Velotype zu verwenden, das in den Niederlanden entwickelt wurde und auf Silbenbasis funktioniert (Romero-Fresco 2011:7). Die BBC begann ihre Live-Untertitelung im Jahr 1990, zuerst mit normalen Tastaturen und dann mit Nutzung von Stenografie. Die Untertitelung übernahmen wie auch in den USA eigens geschulte Stenotypistinnen. Ähnlich war die Entwicklung beim flämischen Sender VRT, der 1984 mit der Live-Untertitelung begann, ebenfalls zuerst mit einer QWERTY-Tastatur, um dann über das Velotype zur Stenografie überzugehen (Geens 2008:101). In anderen Ländern wurden Live-Untertitel erst zu einem wesentlich späteren Zeitpunkt eingeführt, in Spanien und Italien beispielsweise um die Jahrtausendwende.

Grundsätzlich funktionierten die Stenografie-Methoden in beiden Fällen sehr gut und konnten einen hohen Grad an Genauigkeit sowie Geschwindigkeiten von etwa 220 – 250 Wörtern pro Minute erreichen (Romero-Fresco 2011:8, Geens 2008:101). Der problematische Aspekt war allerdings, dass eine Stenotypie-Ausbildung mehrere Jahre in Anspruch nimmt. Daher war diese Methode der Untertitelung sehr teuer. Im Fall von VRT kam noch das Problem hinzu, dass es nicht genügend entsprechend ausgebildete Personen in der Region gab. (Romero-Fresco 2011:8)

Aufgrund der oben genannten Schwierigkeiten bezüglich der Verfügbarkeit der StenotypistInnen und den dadurch anfallenden Kosten wurde bald nach einer Alternative gesucht. Schon in den 1980er-Jahren wurde die Spracherkennung als Zukunftsmöglichkeit in Betracht gezogen, es sollte allerdings noch einige Jahre dauern, bis diese weit genug entwickelt

war, um tatsächlich eingesetzt zu werden und Respeaking als Methode zur Live-Untertitelung zu ermöglichen. Im nächsten Kapitel befasse ich mich mit einer Beschreibung und der Entwicklung der Respeaking-Methode.

2.5 Respeaking

Der Begriff „Respeaking“ und das dazugehörige Forschungsfeld hat in der Translationswissenschaft noch einen vergleichsweise geringen Bekanntheitsgrad. Tatsächlich hat sich der Begriff „Respeaking“ als solcher relativ spät und noch immer nicht vollständig gegenüber anderen Termini durchgesetzt, von welchen Romero-Fresco einige erwähnt. Neben eher unhandlichen erklärenden Begriffen wie *speech recognition based subtitling* gibt es auch kürzere Formen wie zum Beispiel *revoicing* oder *voice-writing* (gebräuchlich in den USA) sowie *shadow speaking* und *speech captioning* (in Australien und Kanada). In Spanien hat sich zudem der sprachlich angepasste Begriff *rehablado* etabliert (Romero-Fresco 2011:2). In dieser Arbeit wird allerdings durchgehend der Begriff Respeaking verwendet. Um aber zu wissen, wovon genau gesprochen wird, ist es notwendig, den Begriff zunächst zu definieren. Ich werde im Folgenden einige Definitionen auflisten, um mich anschließend mit deren Gemeinsamkeiten und Unterschieden zu beschäftigen.

Bemerkenswert ist, dass viele Publikationen zu diesem Thema eine sehr prozessorientierte Beschreibung nutzen, anstatt eine tatsächliche Definition aufzustellen, ähnlich wie jemand die Methode im täglichen Leben einer mit Respeaking nicht bekannten Person beschreiben würde. Ein Beispiel hierfür findet sich in Frederic Chaumes *The turn of audiovisual translation*:

The interpreters use a computer with voice recognition software [...] that they have previously trained to recognise their voice, and listen to the live broadcast through headphones (although this method is also used with recorded broadcasts). The interpreter hears the characters' dialogues or the narrator's voice and re-reads them in his or her own words, usually summarising the original dialogues quite substantially so their re-translation fits into the subtitle space that the software program generates when it processes the spoken sentences. (Chaume 2013:113)

Es handelt sich hierbei aber wie gesagt nicht um eine Definition im eigentlichen Sinne, sondern um eine Beschreibung des Ablaufes. Eine Definition für die Methode, die eher der gängigen Herangehensweise an eine solche entspricht, führt Carlo Eugeni an, der die Technik des Respeakings wie folgt beschreibt:

[...] respeaking is the production of real-time subtitles by means of speech recognition software transcribing a simultaneous reformulation of the source text dictated by the operator, or respeaker, to the computer. (Eugeni 2008:38)

Eine Definition stammt aus dem wohl am häufigsten zitierten Hauptwerk zum Thema Respeaking von Pablo Romero-Fresco. Er beschreibt Respeaking als:

[...] a technique in which a respeaker listens to the original sound of a live programme or event and respeaks it, including punctuation marks and some specific features for the deaf and hard of hearing audience, to a speech recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay. (Romero-Fresco 2011:1)

Beide Definitionen sprechen bereits einige wichtige Merkmale beziehungsweise Problemfelder des Respeakings an, mit denen sich die Literatur in den weiteren Kapiteln noch befassen wird. Diese Aspekte sind die Gleichzeitigkeit des Respeakings – von Eugeni mit „simultaneous“ und von Romero-Fresco mit „shortest possible delay“ gekennzeichnet – und das Hören, Wiederholen und anschließende Verschriftlichen des gesprochenen Textes. Romero-Fresco führt allerdings zwei Dinge an, die Eugeni und auch viele andere Definitionen auslassen. Zum einen erwähnt er die Notwendigkeit, auch Satzzeichen zu respeaken, zum anderen spricht er auch die primäre Zielgruppe der durch Respeaking erzeugten Untertitel an.

2.6 Geschichte und Anwendung des Respeakings

Die Respeaking-Methode wird seit dem Jahr 2001 zur Live-Untertitelung verwendet. Vorreiter bei der Nutzung waren der belgische Fernsehsender VRT und vor allem die britische BBC (Chmiel et al. 2017, Waes et al. 2013, Romero-Fresco 2012). Heute hat sich Respeaking in seiner Verbreitung durchgesetzt und ist in vielen Ländern die Standardmethode zur Live-Untertitelung geworden. Im Vergleich zu anderen Methoden wie zum Beispiel der Verwendung spezieller Tastaturen ist Respeaking zuverlässig und trotzdem kostengünstiger (Waes et al. 2013:17). Im österreichischen Fernsehen wird Respeaking erst seit dem Jahr 2009 eingesetzt (Walter 2012:14).

Obwohl die Methode in jedem Fall auf der Verwendung von Spracherkennung basiert, haben sich in unterschiedlichen Ländern und bei den einzelnen Fernsehsendern verschiedene Herangehensweisen zur Durchführung des Respeakings entwickelt. Hinter diesen Entscheidungen stehen oft Grundsatzfragen darüber, was eine Live-Untertitelung zu leisten hat. Zuerst ergibt sich die Frage der Korrektur der in das Spracherkennungssystem eingelesenen Untertitel. Mit den heutigen technischen Möglichkeiten kann eine perfekte Erkennung durch die Software allein meist noch nicht gewährleistet werden, weshalb oft auf Korrektur durch

Menschen zurückgegriffen wird. In den meisten Fällen korrigiert der/die RespeakerIn selbst, mancherorts, wie zum Beispiel beim französischen Fernsehsender France 2, übernimmt dies aber auch eine zweite Person. Es gibt allerdings auch den selteneren Fall, zum Beispiel bei einem kanadischen Fernsehsender, in dem überhaupt nicht korrigiert wird. Hier wird ein vermehrtes Auftreten von Fehlern akzeptiert, um eine schnellere Ausstrahlung der Untertitel zu ermöglichen (Romero-Fresco 2011:16f).

Diese Debatte bezüglich Geschwindigkeit im Verhältnis zu Korrektheit zieht sich durch die vorhandene Literatur und wird in einem späteren Kapitel noch genauer behandelt. Damit in Verbindung steht auch, wie die Untertitel auf den Fernsehbildschirmen erscheinen. Dies kann entweder im sogenannten *block mode* erfolgen, bei dem eine oder mehrere Zeilen komplett gleichzeitig aufscheinen, oder im *scrolling mode*, wo im Falle von Respeaking der Text Wort für Wort erscheint - bei Methoden, die Tastaturen nutzen, kann dies sogar Buchstabe für Buchstabe geschehen (Luyckx 2013:3). Auch beim Editieren und bei Kürzungen gehen die Meinungen auseinander. Abgesehen von den technischen Möglichkeiten sowie der menschlichen Lesegeschwindigkeit, mit welchen sich ein späteres Kapitel befasst, entscheidet grundsätzlich der jeweilige Sender darüber, ob die RespeakerInnen angehalten sind, den Originaltext nahezu wörtlich oder sinnvoll gekürzt wiederzugeben (Romero-Fresco 2009:111).

Eine weitere Besonderheit von Respeaking ist die durchgehende Verwendung von Software. Dabei kommt Software zur Spracherkennung und zum Formatieren und Senden der Untertitel zum Einsatz. Beispiele für Produkte, die heute zur Untertitelung verwendet werden, sind Dragon Naturally Speaking und Viavoice (Romero-Fresco 2011:17).

3 Forschungsstand zum Thema Respeaking

Die bisherige Forschung zum Thema Respeaking in der Live-Untertitelung konzentriert sich insbesondere auf die in den folgenden Kapiteln beschriebenen Themenbereiche Ausbildung und Kompetenzen, Qualität und Genauigkeit, Geschwindigkeit und Paraphrase sowie die Erwartungen der Zielgruppe und deren Rezeption. Diese Themengebiete werden nach dem theoretischen Teil der Arbeit als Grundlage für die empirische Forschung dienen, weswegen es wichtig ist, sie gründlich aufzuarbeiten.

3.1 Ausbildung und Kompetenzen

Respeaking wird auch als eine Technik, die zwischen audiovisueller Übersetzung und Dolmetschen liegt, beschrieben (Chaume 2013:113). Es scheint daher selbstverständlich, dass in der Literatur häufig auf die ähnlichen Anforderungen an (insbesondere Konferenz-) DolmetscherInnen und RespeakerInnen hingewiesen wird. Dieser Zusammenhang kann auch Einfluss darauf haben, wie RespeakerInnen ausgebildet werden. Momentan findet die Ausbildung jedoch nach wie vor meistens *in-house* bei den jeweiligen Fernsehsendern statt und ist damit sehr variabel und wenig normiert. Ebenso verhält es sich mit dem Rekrutierungsprozess. Eine Ausbildung im Bereich der audiovisuellen Übersetzung wird nur manchmal vorausgesetzt, eher aber ein Diplom in einer zeitgenössischen Sprache (Romero-Fresco 2012:97f).

In Großbritannien zum Beispiel beinhaltet der Aufnahmeprozess von RespeakerInnen einen Aufnahmetest, in dem eine Sendung mit einigen üblichen Schwierigkeiten – zum Beispiel sich überschneidenden Dialogen – durch Respeaking untertitelt werden soll. Wird ein/e KandidatIn eingestellt, folgt eine Einschulung, die je nach den Fähigkeiten des/der KandidatIn zwischen drei Wochen und drei Monaten andauert. Als Fähigkeiten, auf die bei diesem Prozess geachtet wird, werden eine klare, ausgeglichene Sprache, gleichbleibende Vortragsweise, sowie ein gutes Verständnis des Ausgangstextes und des selbst produzierten Textes genannt (Romero-Fresco 2011:24). In Spanien wird von den KandidatInnen keine besondere Vorbildung erwartet, sie sollen allerdings exzellente Sprach-, Editier- und Teamwork-Fähigkeiten haben und sich außerdem für die Probleme von gehörlosen und schwerhörigen Personen beim Konsum von Fernsehen interessieren. In Flandern wird nach zwei bis drei Wochen Training entschieden, ob ein/e UntertitlerIn auch in der Live-Untertitelung arbeiten kann. Sollte dies der Fall sein, dauert die Ausbildung noch weitere drei Monate (Romero-Fresco 2011:27f).

Wie sich zeigt, werden die Fähigkeiten eines Respeakers/einer RespeakerIn als wichtiger als die vorhergehende Ausbildung erachtet. Darunter fallen im Allgemeinen exzellente Sprachkenntnisse, insbesondere in den Bereichen Interpunktion und Rechtschreibung und gute Kommunikation. Editierfähigkeit und das Arbeiten im Team werden ebenso als wichtig erachtet. Auch sollen angehende RespeakerInnen gut über das aktuelle Weltgeschehen informiert sein. Ein weiterer wichtiger Faktor ist die Fähigkeit zum Multitasking, da bei der Live-Untertitelung durch Respeaking mehrere Aufgaben zugleich erfüllt werden müssen (Romero-Fresco 2011). Laut Romero-Fresco scheint es bei den Fernsehsendern wenig Bewusstsein über Kurse in Respeaking im universitären Bereich zu geben (Romero-Fresco 2012:97f). Es gibt jedoch schon einige Sender, wie zum Beispiel den belgischen VRT, die die Zusammenhänge zwischen Dolmetschen und Respeaking erkannt haben. Dieser Fernsehsender beauftragte Rita Geens, eine Lehrende an der Universität Gent, die dort arbeitenden RespeakerInnen in Dolmetschtechniken zu unterrichten (Geens 2008:105).

Obwohl die Ausbildung wie oben beschrieben normalerweise bei den Fernsehsendern direkt erfolgt, werden auch Anstrengungen unternommen, Untertitelung für Gehörlose und damit auch Respeaking in die Dolmetsch-Curricula einzugliedern. Diese Entwicklung begann im Jahr 2008. In Spanien war die Universität, die hier als Vorreiter agierte, die Universität Autònoma de Barcelona im Rahmen des Masterprogrammes *Audiovisual Translation* (MTAV). Das Modul *Superimposed Voices and Respeaking* umfasst momentan 6 ECTS (Universität Autònoma de Barcelona 2020). In Großbritannien bietet die University of Roehampton im MA *Audiovisual Translation* innerhalb des Moduls *Audiodescription, Subtitling for the Deaf and Respeaking* Kurse im Respeaking an. Sie waren im Jahr 2012 die erste und einzige Universität, die Respeaking-Kurs auch in verschiedenen Sprachen anbot (University of Roehampton 2020). Seit 2008 werden Respeaking-Kurse auch an der Universität Gent angeboten (Geens 2008).

Das Artesis University College in Antwerpen war hingegen die erste Institution, die RespeakerInnen im Rahmen eines Masterkurses im Dolmetschen ausbildete. Auch in Österreich werden universitäre Kurse auf diesem Gebiet angeboten. An der Universität Wien wird seit dem Wintersemester 2018 ein 4 ECTS-Kurs zur Live-Untertitelung mittels Respeaking angeboten, der auch interlinguales Untertiteln von Englisch nach Deutsch beinhaltet (Universität Wien 2018). An der Theoretischen und Angewandten Translationswissenschaft an der Universität Graz wird ein ebenfalls 4 ECTS umfassender Kurs zu *Schriftdolmetschen und Respeaking* unterrichtet (Karl-Franzens-Universität Graz 2017:7).

3.1.1 Gemeinsamkeiten zwischen Respeaking und Dolmetschen

Die auffälligste Gemeinsamkeit zwischen Respeaking und dem simultanen Dolmetschen ist die Gleichzeitigkeit des Zuhörens und Sprechens (Arumi-Ribas & Romero-Fresco 2008:110). Laien könnten aufgrund der Intralingualität von Respeaking fälschlicherweise annehmen, ein/e RespeakerIn würde lediglich den Ausgangstext exakt wiederholen. In Wirklichkeit sind RespeakerInnen aber häufig gezwungen, den Text zu adaptieren. Dies kann verschiedene Gründe haben, etwa die Notwendigkeit zur Verkürzung aus Zeitmangel, das Entfernen beziehungsweise Korrigieren von verwirrenden Versprechern oder Fehlstarts, oder die Vermeidung von Worten, die nicht im Vokabular der Spracherkennungssoftware vorhanden sind und daher von dieser wahrscheinlich nicht erkannt werden. Obwohl ein/e RespeakerIn also nicht in eine andere Sprache transferiert, sind Fähigkeiten im Bereich der Sinnerfassung und der Umformulierung trotzdem sehr wichtig (Geens 2008:103). Szarkowska et al. erwähnen in einer 2016 veröffentlichten Arbeit ebenfalls die für das Respeaking notwendige Fähigkeit gleichzeitig zu hören und zu sprechen, und untersuchen die damit einhergehende hohe kognitive Belastung. Sie verglichen außerdem die Höhe dieser kognitiven Belastung zwischen DolmetscherInnen, zweisprachigen Personen und ÜbersetzerInnen. Hierbei kamen sie zu dem Schluss, dass DolmetscherInnen tatsächlich in einigen Kategorien eine geringere Belastung durch das Respeaking aufwiesen (Szarkowska et al. 2016).

Eine weitere mit der Gleichzeitigkeit in Zusammenhang stehende Gemeinsamkeit ist die Live-Situation, in der beide Handlungen ausgeführt werden und die wenig Raum für nachträgliche Korrekturen bietet. Weitere Gemeinsamkeiten betreffen die Vorbereitung sowie die Umgebung, in der das Respeaking stattfindet. Eine Vorbereitung durch Glossare ist in beiden Fällen wichtig, auch wenn sie beim simultanen Dolmetschen für den/die DolmetscherIn selbst erfolgt und beim Respeaking dazu dient, das Spracherkennungsprogramm auf neue, noch nicht im Vokabular enthaltene Begriffe zu trainieren. Auch arbeiten sowohl RespeakerInnen als auch DolmetscherInnen mit Kopfhörern und Mikrofonen, und in manchen Fällen findet Respeaking auch in Kabinen statt, obwohl dies nicht zwangsläufig so sein muss.

Außerdem finden beide Arbeiten in einem Team statt, da DolmetscherInnen und RespeakerInnen nach einer gewissen Zeit abgelöst werden müssen. In beiden Fällen soll der Zeitabstand zwischen dem Ausgangs- und Zieltext so gering wie möglich sein und es ist für beide Tätigkeiten wichtig, die eigene Stimme kontrollieren zu können. Die Anforderungen bezüglich der Stimme sind allerdings unterschiedlicher Natur. Für eine/n DolmetscherIn ist es wichtig, für die Zuhörer angenehm zu sprechen und auch Sprachfärbungen wie Emotionen

wiederzugeben. Hingegen sind RespeakerInnen dazu angehalten, eher monoton und ohne besondere Färbung zu sprechen, da dies die Erkennung für die Software erleichtert (Arumi Ribas & Romero Fresco 2008: 110).

Ein wichtiger Unterschied ist das fertige Produkt. Bei einer Dolmetschung wird ein gesprochener Input in einen gesprochenen Output umgewandelt. Obwohl die Arbeit eines Respeakers/einer Respeakerin auch mündlich erfolgt, erstellt diese/r nur eine Art Zwischentext, welcher dann von der Spracherkennungssoftware in ein schriftliches Endprodukt umgewandelt wird. Das führt dazu, dass beim Respeaking auch Aspekte der schriftlichen Sprache wie Interpunktion miteinbezogen werden müssen. Weiters ist eine Dolmetschung in jedem Fall interlingual, während Respeaking oft intralingual erfolgt und sich auch diese Arbeit mit der intralingualen Variante befasst (Arumi Ribas & Romero Fresco 2008:111).

Arumi Ribas und Romero-Fresco stellen anhand dieser Unterschiede und Gemeinsamkeiten eine Kompetenzen-Matrix auf. Neben den bereits oben erwähnten Faktoren kristallisiert sich hier heraus, dass die dem Respeaking eigenen Fähigkeiten besonders mit der Nutzung der Technik und dem Umgang mit einem weiteren gleichzeitigen Input in Form des Textes am Bildschirm zusammenhängen. Außerdem kommen gewisse Merkmale zu tragen, die der Untertitelung als solcher zu eigen sind (zum Beispiel die unterschiedliche Färbung des Textes je nach Sprecher), zu tun haben. Arumi Ribas und Romero-Fresco schlagen verschiedene Trainingsmethoden für RespeakerInnen vor, um diese Kenntnisse zu erlangen. (Arumi Ribas & Romero-Fresco 2008:110f, Romero-Fresco 2011: 46f)

Aus der Ähnlichkeit von Respeaking und Simultandolmetschen ergeben sich bezüglich der Ausbildung von RespeakerInnen zwei Überlegungen: Erstens zur Frage, wie man ausgebildete SimultandolmetscherInnen zu RespeakerInnen weiterbilden kann, und auf der anderen Seite, wie man die Ausbildung von RespeakerInnen durch Techniken der Dolmetschdidaktik erweitern kann. Rita Geens, die wie oben erwähnt vom Sender VRT beauftragt wurde, RespeakerInnen in Dolmetschtechniken zu unterrichten, erwähnt hierzu keine Details, widmet in ihrer Arbeit zum Thema Respeaking aber einige kurze Absätze der Thematik der Ausbildung von KonferenzdolmetscherInnen zu RespeakerInnen. Sie beschreibt ein Training zur Reduktion des Textes durch Auslassungen und Umformulierungen, um ihn für die kurzen Untertitel passend zu machen. Außerdem beinhaltet diese Ausbildung die Segmentierung. Untertitel sollten komplexe Sätze zu Gunsten von kürzeren, einfacheren Sätzen vermeiden (Geens 2008:106ff).

Arumi Ribas und Romero-Fresco beschreiben hingegen konkretere Übungen. Die von ihnen genannten Themenbereiche sind die Vorbereitung inklusive des richtigen Umgangs mit

der Spracherkennungssoftware und deren Vokabeldatenbank, die Fähigkeit zum Hörverständnis und zur Analyse von Texten, die Fähigkeit zum Umformulieren, Multitasking, Umgang mit einer Live-Situation sowie die richtige Vortragsweise. Die Übungen sind dabei sehr praxisorientiert. Um die Fähigkeit zur Umformulierung zu stärken, sollen sich die TeilnehmerInnen beispielsweise eine Rede anhören und diese dann in ihren eigenen Worten wiedergeben, möglicherweise auch in einem anderen Register als der Originalsprecher oder mit blumigeren Worten (Arumi Ribas & Romero-Fresco 2008).

In dieser Debatte um die Ausbildung von DolmetscherInnen zu RespeakerInnen und den Ähnlichkeiten zwischen den Anforderungen beider Disziplinen sollte aber nicht vergessen werden, dass die Annahme, dass sich Personen mit Simultandolmetsch-Ausbildung automatisch besser als RespeakerInnen eignen, keineswegs bewiesen ist. In einer 2017 herausgegebenen Studie testete ein Team von polnischen WissenschaftlerInnen die Fähigkeiten zum Umformulieren (Paraphrasieren) und die Pausen beziehungsweise Verzögerung beim Respeaking in einem Respeaking-Experiment. An dem Experiment nahmen SimultandolmetscherInnen, ÜbersetzerInnen und eine Kontrollgruppe aus zweisprachigen Personen teil. Alle TeilnehmerInnen bekamen in einem Zeitraum von nur zwei Tagen eine Art Crashkurs im Respeaking. Es konnten bei dem darauffolgenden Experiment je nach Themengebiet und Textart sehr geringe bis gar keine Unterschiede zwischen der Leistung der DolmetscherInnen und jener der beiden anderen Gruppen festgestellt werden. Natürlich handelt es sich hierbei um eine sehr eingeschränkte Personengruppe, die Ergebnisse sind allerdings doch bemerkenswert und zeigen, dass zwischen Simultandolmetschen und Respeaking trotz der offensichtlichen Gemeinsamkeiten noch immer ein großer Unterschied besteht, der sich nicht so einfach überbrücken lässt (Chmiel et al. 2017).

3.1.2 Hintergründe der RespeakerInnen

Robert, Shrijver und Diels von der Universität Antwerpen veröffentlichten im Jahr 2019 eine Studie zu den Hintergründen von RespeakerInnen in 27 verschiedenen Ländern. Ziel der Forschung war es, ein Kompetenzenprofil für die interlinguale Untertitelung zu erstellen. Da aber die meisten RespeakerInnen (auch) intralingual arbeiten, waren die Befragten auch als intralinguale UntertitlerInnen tätig, wodurch die Arbeit auch für diese Forschung Relevanz hat. Bezüglich des akademischen Hintergrundes der TeilnehmerInnen ergab die Umfrage, dass der Großteil über eine abgeschlossene universitäre Ausbildung verfügte. Bei den Studienrichtungen waren Sprachstudien vorherrschend, mit einem großen Anteil an

TranslationswissenschaftlerInnen. Die meisten arbeiteten nicht nur als Live-UntertitlerInnen, sondern auch in anderen Bereichen wie beispielsweise als ÜbersetzerInnen oder UntertitlerInnen im nicht-live Bereich. 76 Prozent der intralingualen UntertitlerInnen wurden *in-house* ausgebildet. Andere Formen der Ausbildung, wie zum Beispiel jene im universitären Bereich, nehmen also weiterhin eine geringere Rolle ein. Die Ausbildung war in den meisten Fällen praktisch orientiert und befasste sich mit dem Umgang mit der Software sowie Diktat- und Respeakingübungen. In manchen Fällen gab es gar keine dezidierte Ausbildung, sondern es wurde bei der Arbeit selbst gelernt. Die meisten TeilnehmerInnen bewerteten ihre Ausbildung sehr positiv, auch wenn sie sich in gewissen Bereichen noch detailliertere oder längere Einführungen gewünscht hätten (Robert et al. 2019).

3.2 Qualität und Genauigkeit

Um die Qualität einer durch Respeaking erstellten Untertitelung evaluieren zu können, muss es diesbezüglich natürlich Kriterien geben. Die häufigste Methode, die Qualität zu quantifizieren, ist eine Form der Fehlerberechnung. Diese kann entweder von RespeakerInnen gegenseitig nach dem Einsatz durchgeführt werden oder auch von TrainerInnen. Das bedeutet allerdings nicht, dass diese Fehlerraten das einzige Qualitätsmerkmal einer Live-Untertitelung sein sollten, andere qualitative Kriterien wie die Rezeption durch die Zielgruppen, Verständlichkeit und Ähnliches sollten ebenfalls miteinbezogen werden. Während eine mathematisch basierte Methode des Berechnens von Fehlern eine schnelle Möglichkeit ist, sich einen ersten Überblick zu verschaffen, so muss sie für eine sinnvolle Nutzung für das Respeaking adaptiert werden. Im Bereich der Spracherkennung wird normalerweise das WER-Modell (*Word Error Rate Model*) verwendet. Dieses Modell wertet aber generell alle Abweichungen vom oder Auslassungen gegenüber dem Ausgangstext automatisch als Fehler. Da aber ein/e RespeakerIn angehalten ist, aus verschiedenen Gründen den Ausgangstext anzupassen, ist eine Wertung als Fehler nicht immer gerechtfertigt (Romero-Fresco 2011:151).

Laut Romero-Fresco muss eine Berechnungsmethode die folgenden Kriterien erfüllen: Sie muss funktional und einfach anzuwenden sein, die bewährte WER-Berechnungsmethode einschließen, unterschiedliche Kriterien für verschiedene Fernsehprogramme miteinbeziehen, und eine Möglichkeit bieten, die verschiedenen Standards in den einzelnen Ländern zu berücksichtigen. Außerdem soll im durch Respeaking produzierten Text zwischen Editier- und Erkennungsfehlern unterschieden und diese nach ihren Schweregrad eingeschätzt werden. Wenn möglich, sollten auch andere relevante Informationen wie Verzögerung, Position,

Geschwindigkeit o.ä. miteinbezogen werden. Wichtig ist auch, dass das Modell Anregungen dafür gibt, welche Aspekte verbessert werden sollen und möglicherweise sogar, wie dies geschehen soll (Romero-Fresco 2011:150). Trotz ihrer Schwächen im Bereich des Respeakings werden Modelle wie das WER-Modell häufig angewandt, daher folgt nun eine Beschreibung ihrer Entwicklung.

3.2.1 Die Wortfehlermodelle

Das *US National Institute of Standards and Technology* misst die Wortgenauigkeit mit folgender Formel:

$$\text{Accuracy rate} = \frac{N - \text{Errors}}{N} \times 100 = \%$$

Abb. 1: Das WER-Modell (Romero-Fresco & Martinez Perez 2015:30)

N bezeichnet hierbei die Anzahl der Wörter, die gesprochen wurden. Aufbauend auf diese sehr einfache Formel stellt das Institut auch noch eine Berechnungsweise für Wortkorrektheit auf. Diese unterscheidet zwischen drei verschiedenen Fehlertypen, die beim Respeaking auftreten können, nämlich Löschung (ein Wort wird ausgelassen – in der Formel als D angegeben), Ersetzung (ein falsches Wort wird statt des korrekten eingefügt – in der Formel als S angegeben) und Einfügen (ein Wort wird eingefügt, wo vorher keines stand – in der Formel als I angegeben). Die daraus entstehende Berechnung sieht wie folgt aus, die Zahlen sind als Beispiel gewählt:

$$\text{Accuracy rate} = \frac{N(6) - D(1) - S(3) - I(1)}{N(6)} \times 100 = 16.66\%$$

Abb. 2 WER-Modell II (Romero-Fresco & Martinez Perez 2015:30)

Diese Methode unterscheidet aber nicht, ob eine Auslassung oder Ersetzung die Bedeutung verändert hat oder nicht.

Der erste Versuch, diese Art der Berechnung an die Realität des Respeakings anzupassen, war die CRIM-Methode, erarbeitet durch das *Centre de Recherche Informatique de Montréal*, bei der nach der Berechnung von Abweichungen des Textes noch eine menschliche Kontrolle erfolgt, ob diese den Sinn verändern (Romero-Fresco & Martinez Perez

2015, Romero-Fresco 2011). Eine Weiterentwicklung dieser Methodik ist das sogenannte NERD-Modell von Romero-Fresco. Zur Berechnung der Genauigkeit wird hierbei N als die Anzahl der Worte im Respeaking-Text definiert. Davon abgezogen werden die Editierfehler, die durch eine Entscheidung des/der RespeakerIn entstehen, sowie Erkennungsfehler, die von einer Fehlerkennung der Spracherkennungssoftware stammen. Ebenfalls abgezogen werden die sogenannten *deducted marks*, wobei für korrigierte Fehler weniger abgezogen wird, beziehungsweise können in manchen Fällen auch weniger schwere Fehler hierunterfallen – Romero-Fresco nennt als Beispiel hierfür Akzentfehler im Spanischen. Außerdem gibt es die Möglichkeit, korrekte Veränderungen – also solche, die die Bedeutung nicht verändert haben – sowie schwere Fehler gesondert anzuführen. Diese sind aber nicht in der Formel enthalten, da es sich bei ersteren nicht um Fehler handelt und zweitens zu schwierig zu kategorisieren sind, da die Kriterien je nach Art der Sendung variieren. Auch wird eine Möglichkeit angeboten, Kommentare bezüglich anderer Qualitätsmerkmale des Respeakings zu machen. Romero-Fresco sieht bei dieser Methode eine Genauigkeit über 97 bis 97,5% als akzeptabel an. (Romero-Fresco 2011:153f).

In einer Veröffentlichung aus dem Jahr 2015 haben Romero-Fresco und Martinez Perez das NERD-Modell zu einem neuen NER-Modell weiterentwickelt. Der größte Unterschied ist, dass hier die von Romero-Fresco als *loose category* bezeichneten *deduction marks* fehlen. Auch gibt es keine gesonderte Anführung von schweren Fehlern mehr. Stattdessen werden schwere, mittlere und leichte Fehler bereits in die Berechnung einbezogen und jeweils mit 1, 0,5 und 0,25 gewichtet. Die Berechnung sieht damit wie folgt aus:

$$Accuracy = \frac{N - E - R}{N} \times 100$$

CE (correct editings):
Assessment:

Abb. 3 NER-Model (Romero-Fresco 2015:32)

Noch immer findet sich die Möglichkeit, die korrekten Editierungen und eine Bewertung, die wie die im NERD-Modell erwähnten Kommentare dazu dienen soll, Informationen außerhalb von Worten und Fehleranzahl zu geben, anzugeben (Romero-Fresco 2015:32f). Nach der Analyse dieses Modells muss das Respeaking laut Romero-Fresco eine Genauigkeit von 98% erreichen, um als qualitativ in Ordnung zu gelten.

Die Qualitätsstandards und die Methoden, diese Qualität zu bewerten, variieren stark zwischen und teilweise sogar innerhalb der einzelnen Länder. Daher ist es relativ schwierig, hier einen Vergleich zu ziehen. Das oben beschriebene NER-Modell wird in Untertitelungsfirmen, Rundfunkanstalten, Universitäten und Regierungsorganisationen in mehr als 13 Ländern verwendet, darunter in Großbritannien, Spanien, den DACH-Staaten, aber auch Südafrika, Kanada und Australien. Außerdem wird es von der ITU Focus Group on Media Accessibility der Vereinten Nationen unterstützt. In den USA hingegen wird die Qualität der Untertitelung momentan nicht mit einem quantitativen Modell überprüft, stattdessen werden Veränderungen auf Basis von individuellen Zuseherbeschwerden vorgenommen (Romero-Fresco 2016:60).

3.2.2 Guidelines für Qualität

Wie bereits erwähnt, gibt es aber auch noch andere Qualitätsmerkmale für Live-Untertitelung abseits von Fehlerraten und Genauigkeits-Prozenten. Auch hierzu geben einige Länder Qualitätsleitfäden für ihre jeweiligen Rundfunkanstalten heraus. Neben der bereits besprochenen Genauigkeit spielt vor allem der Geschwindigkeitsfaktor, bestehend aus der Zeichenanzahl der Untertitel pro Minute und der Verzögerung, mit der diese in Verhältnis zum Ausgangstext gesendet werden, eine Rolle. Die Geschwindigkeit wird in einem weiteren Unterkapitel noch näher erläutert werden.

In Europa haben drei Länder solche Qualitätsleitfäden aufgestellt: Frankreich, Spanien und Großbritannien. In Frankreich gibt der *Conseil Supérieur de l'Audiovisuel* (CSA) Richtlinien für Länge, Platzierung, Präsentation, Lesbarkeit, Segmentierung, Editierung und Geschwindigkeit und Verzögerung heraus. In Spanien werden, neben der Geschwindigkeit, die Präsentation, Rechtschreibung, Typographie, Sprecheridentifikation und die Übertragung von Soundeffekten als Qualitätsmerkmale genannt. Ofcom hat in Großbritannien vier Merkmale einer guten Untertitelung festgelegt: Genauigkeit, Geschwindigkeit, Verzögerung und Präsentation (Fresno 2018).

Hierbei handelt es sich natürlich nur um Beispiele, und eine solche Vorgabe kann auch nicht als Absicherung gesehen werden, dass diese Standards wirklich immer erreicht werden. Lediglich soll gezeigt werden, dass die Faktoren Genauigkeit, Präsentation in verschiedener Form sowie die Geschwindigkeit der Untertitel durchwegs als wichtige Qualitätsmerkmale gesehen werden, auch wenn sie nicht in allen Fällen einen genauen numerischen Wert zugewiesen bekommen.

3.2.3 Geschwindigkeit

In der Diskussion um die Qualität der Untertitel spielt die Geschwindigkeit eine große Rolle. Beim Thema Geschwindigkeit geht es natürlich um technische Fragen, allerdings ist sie nicht nur ein technisches Problem, sondern auch ein ökonomisches und politisches, bei dem die Interessen von Sendern und der Zielgruppe oft aufeinanderprallen (Romero-Fresco 2009:110).

Grundsätzlich kann eine Person Gehörtes wesentlich schneller verarbeiten als Gelesenes. Daher kann eine Textgeschwindigkeit, die für Hörende akustisch einwandfrei verständlich ist, als Untertitel bereits zu schnell sein. Das Problem wird noch dadurch verstärkt, dass im Vergleich zu spontaner Alltagssprache zuvor eingeübte und abgelesene Fernsehtexte häufig noch schneller gesprochen sind. Außerdem machen Fernsehsprecher allgemein auch weniger häufig Pausen, was die Geschwindigkeit zusätzlich erhöht (Romero-Fresco 2009:113).

Zum Thema der Lesegeschwindigkeit wurden verschiedene Studien durchgeführt, mit sehr unterschiedlichen Ergebnissen, die von einer Geschwindigkeit von 315 Wörtern pro Minute bis hinunter zu nur 200 reichen. Diese Geschwindigkeiten beziehen sich allerdings auf gedruckten Text, bei Untertiteln scheint die Lesegeschwindigkeit grundsätzlich langsamer zu sein. In der Untertitelung haben sich nach Untersuchungen von D'Ydewalle, die von Diaz Cintas bestätigt wurden, 144 Wörter pro Minute (oder eine Standzeit von 6 Sekunden für zwei Standard-Untertitelzeilen) als ideale Geschwindigkeit für die Untertitelung etabliert. Bei diesen Untersuchungen handelt es sich jedoch um hörende Probanden, nicht um hörbehinderte Personen. Darüber hinaus ist die Lesegeschwindigkeit von verschiedenen Faktoren, wie beispielsweise dem Bildungsgrad, abhängig und kann daher sehr schwer generalisiert werden. Bei hörbehinderten Personen sind diese Unterschiede aufgrund der inhärenten Inhomogenität der Gruppe mindestens ebenso gravierend, und es ist schwierig, eine allgemeine Aussage über Lesegeschwindigkeiten zu treffen.

Es gibt bis jetzt wenig Forschung dazu, wie schnell durch Respeaking produzierte Untertitel im Allgemeinen sind. Studien von Eugeni und Romero-Fresco ergaben 2009 allerdings, dass RespeakerInnen grundsätzlich langsamer sind als ihr Ausgangstext, und zwar sogar dann, wenn der Ausgangstext für Fernsehverhältnisse schon in relativ geringer Geschwindigkeit gesendet wird, wie beispielsweise bei Sportsendungen. Die tatsächliche Geschwindigkeit in ihren Untersuchungen bei der BBC lagen zwischen 106 und 190 Wörtern pro Minute. Diese Studie wird im Kapitel zum Thema Wörtlichkeit und Paraphrase noch weiter

aufgearbeitet werden, da es sich hierbei um verwandte Themebereiche handelt (Romero Fresco 2010:176f).

Ein weiterer wichtiger Aspekt der Geschwindigkeit ist die Verzögerung zwischen Ton und Untertitelung. Aufgrund des Untertitelungsprozesses gibt es diese Verzögerung in jedem Fall, auch wenn die Dauer variieren kann. Auch bezüglich der Geschwindigkeit und Verzögerung haben die einzelnen Länder Richtlinien herausgegeben. In Frankreich sollen die Untertitel um nicht mehr als 10 Sekunden verzögert sein, während in Spanien nicht mehr als 15 Zeichen pro Sekunde mit nicht mehr als 8 Sekunden Verzögerung gesendet werden. Laut dem UK Office of Communications (Ofcom) liegt die maximale Geschwindigkeit bei 200 Wörtern pro Minute und die Verzögerung sollte nicht mehr als 3 Sekunden betragen. Diese kurze Verzögerung wird durch die Nutzung von *scrolling mode* statt *block mode* erreicht (siehe Kapitel 2.5)

3.3 Wörtlichkeit oder Paraphrase

Abgesehen von numerischen Aspekten wie Verzögerung oder die Anzahl an Fehlern in einem gegebenen Text kann man jedoch die Qualität von Untertitelung auch noch auf tiefergehender Ebene besprechen. Tatsächlich besteht, grundsätzlich betrachtet, wenig Einigkeit darüber, was gute Untertitelung ausmacht. Ein in diesem Zusammenhang debattiertes Thema ist, wie wörtlich Untertitelung sein sollte. Hier spielt neben technischer Machbarkeit auch die Präferenz der Zielgruppen eine Rolle. Bevor man sich jedoch im Zusammenhang mit Respeaking mit Wörtlichkeit oder Paraphrase befassen kann, muss erst einmal festgestellt werden, was Paraphrase eigentlich bedeutet. Eine Definition aus der Literatur vergleicht den Prozess des Paraphrasierens mit jenem der Translation:

Paraphrasing can be seen as intralingual translation or ‘rewording’ (Jakobson, 1959), where the same language is used to convey the same message using different phrasing. (Chmiel 2017:3)

Wie bereits erwähnt, unterscheidet man zwei Arten, wie Untertitel im Fernsehen ausgestrahlt werden können. Auf der einen Seite der sogenannte *block mode*, bei dem Untertitel als ganzer Block zugleich auf dem Bildschirm erscheinen, auf der anderen Seite der *scrolling mode*, bei dem die Untertitel Wort für Wort erscheinen. Diese Unterscheidung ist auch eng verbunden mit editierten beziehungsweise paraphrasierten und wörtlichen Untertiteln. Erstere sind beim *block mode* und zweitere beim *scrolling mode* vorherrschend. Der *block mode* setzt wegen der begrenzten Lesegeschwindigkeit und dadurch eingeschränkten Anzahl an Zeichen eine gewisse

Verkürzung voraus. Hingegen kann beim *scrolling mode* eine größere Wörtlichkeit erreicht sein, dafür ist hier eine hohe Lesegeschwindigkeit Voraussetzung, dieser Geschwindigkeit auch adäquat folgen zu können (Remael 2007:31).

Die Frage, ob Untertitelung in wörtlicher oder editierter Form erfolgen soll, ist umstritten. Während viele Forschende die Ansicht vertreten, dass editierte Untertitel eine bessere Wiedergabe des Gesprochenen erlauben, fordern Gehörlosenorganisationen im Sinne der Vollständigkeit oft eine wörtliche Wiedergabe. Diese Forderung gilt nicht nur für vorgefertigte Untertitel, sondern auch für Live-Untertitelung. Die Entscheidung zwischen editierten und wörtlichen Untertiteln steht auch in Zusammenhang mit dem Qualitätsmerkmal der Geschwindigkeit, in dem Sinne, dass editierte Untertitel normalerweise langsamer sind als wörtliche (Romero-Fresco 2009:109f).

Auf die Thematik der Forderungen der Hauptzielgruppe der Gehörlosen und hörbeeinträchtigten Personen wird im nächsten Kapitel noch näher eingegangen. Abgesehen von diesen Erwartungen, ist es allerdings auch sinnvoll, sich damit zu beschäftigen, wie realistisch ihre Umsetzung ist. Denn selbst wenn man davon ausginge, dass wörtliche Untertitel eine bessere Informationswiedergabe ermöglichen, bleibt immer noch die Frage, inwiefern diese auf dem momentanen Stand der Technik überhaupt möglich sind. Bisher gibt es wie in vielen Bereichen des Respeakings noch kaum Forschung zu dem Thema.

Im Jahr 2009 führte Romero-Fresco eine Untersuchung zu diesem Thema bei dem britischen Fernsehsender BBC durch. Die BBC ist insofern ein gutes Forschungsobjekt für diese Thematik, da die RespeakerInnen dazu angehalten sind, möglichst wörtlich wiederzugeben (Romero-Fresco 2009:120). Für die Studie wurden zehn jeweils fünf Minuten lange Ausschnitte der im Fernsehen ausgestrahlten Untertitel aus verschiedenen Sendungsgenres analysiert. Die Geschwindigkeit des Originaltons reichte hierbei von einer relativ langsamen Geschwindigkeit von 124 wpm (Wörter pro Minute) bis hin zu 232 wpm. Interessant ist allerdings, dass die durch Respeaking produzierten Untertitel in allen Fällen langsamer waren als der Originaltext. Das bedeutet, dass obwohl die RespeakerInnen im Fall der schnellsten gesprochenen Originaltexte in einer Geschwindigkeit von bis zu 190 wpm Untertitel produzieren konnten, sie dennoch auch bei der Geschwindigkeit von nur 124 wpm langsamer als die OriginalsprecherInnen waren. Die Begründung dafür liegt zunächst in der Notwendigkeit von Satzzeichen. Tatsächlich sprachen die RespeakerInnen zumindest bei den langsameren Segmenten in einer ähnlichen Geschwindigkeit wie der Ausgangstext, da aber die Satzzeichen in ihrem Fall auch in die Anzahl an gesprochenen Worten fallen, war der Output insgesamt langsamer. Daraus ergibt sich laut Romero-Fresco, dass die RespeakerInnen im

Durchschnitt sogar 20 bis 25 Wörter pro Minute schneller sprechen müssten um dieselbe Output-Geschwindigkeit wie ihr Ausgangstext zu erhalten und somit Wörtlichkeit zu gewährleisten (Romero-Fresco 2009: 122f). Es reicht also nicht aus, wenn RespeakerInnen in derselben Geschwindigkeit wie der Ausgangstext wiedergeben können, was eine tatsächlich wörtliche Wiedergabe in noch weitere Ferne rückt.

3.3.1 Informationsverlust durch Reduktion?

Da häufig von der Gefahr eines Informationsverlusts durch editierte beziehungsweise nicht-wörtliche Untertitel gesprochen wird, lohnt es sich außerdem sich damit zu befassen, inwiefern ein solcher Verlust von Information überhaupt stattfindet, und zu welchen Inhalten der/die ZuhörerIn dadurch tatsächlich keinen Zugang mehr hat. Um dies zu untersuchen nutzt Romero-Fresco in seiner BBC-Studie die Unterteilung in sogenannte *idea units*. Diese müssen folgende vier Kriterien erfüllen:

- They are spoken with a single coherent intonation contour, ending in what is perceived as a clause-final intonation;
- They are preceded and followed by some kind of hesitation, ranging from a momentary break in timing to a filled or unfilled pause lasting several seconds;
- They are clauses -that is, they contain one verb phrase along with whatever noun phrases, prepositional phrases, adverbs, and so on are appropriate;
- They are about seven words long and take about two seconds to produce.

(Romero-Fresco 2009:125)

Durch das Identifizieren und Analysieren dieser *idea units* im durch Respeaking produzierten Text und einem Vergleich mit dem Originaltext versucht Romero-Fresco festzustellen, wie viel Information tatsächlich durch die Reduktion des Textes verloren gegangen ist. Ergebnis der Untersuchung war, dass die Reduktion der *idea units* verglichen mit der Reduktion des Gesamttextes prozentuell gesehen durchgehend geringer war. Das bedeutet, dass trotz der Textreduktion eine hohe Anzahl an *idea units* und somit Information erhalten blieb (Romero-Fresco 2009:126).

Dies zeigt auch, welche Arten von Äußerungen im Allgemeinen durch RespeakerInnen reduziert werden. In einer qualitativen Analyse identifiziert Romero-Fresco die meisten Auslassungen als einem der folgenden vier Typen von Äußerungen angehörend: Diskursmarker wie „so, well, I mean“, Konnektoren wie „and“ oder „but“, Verstärkungswörter wie

„really“ und „very“, oder Wiederholung beziehungsweise unwichtige Worte wie „worth bearing in mind“, „as you were saying“. (Romero-Fresco 2009:127)

In einem in Polen durchgeführten Experiment wurde eine Gruppe von TeilnehmerInnen angehalten, einen Text zu paraphrasieren und als Respeaking wiederzugeben. Ihnen wurde als Aufgabe gestellt, die Länge der jeweiligen Sätze um etwa 30 Prozent zu reduzieren. Diese Sätze wurden dann auf ihre semantische Äquivalenz sowie darauf, ob als redundant oder unnötig angesehene Satzteile wirklich entfernt wurden, analysiert. Während die Zielsetzung der Studie darin bestand, Unterschiede zwischen den Leistungen von zweisprachigen Personen, ausgebildeten ÜbersetzerInnen und DolmetscherInnen zu finden, ist die Analyse der semantischen Äquivalenz in diesem Zusammenhang interessant, da sie zeigt, welche Informationen für die ForscherInnen als paraphrasierbar und damit für den Informationsgehalt als überflüssig erachtet werden. Dabei handelt es sich um eben jene Äußerungsarten, die oben genannt wurden, wie Diskursmarker, Wiederholungen und die sogenannten *false starts* (Chmiel 2017:17f).

Auch wenn in den Studien durch diese Auslassungen kein Informationsverlust festgestellt wird beziehungsweise nicht davon ausgegangen wurde, dass ein solcher besteht, ist dies natürlich schwierig zu quantifizieren. Geht man aber davon aus, dass schwerhörige und gehörlose Personen eine Sendung ebenso erleben sollen wie Hörende, ist es fraglich, ob eine Wiederholung oder ein Verstärkungswort hierbei nicht doch eine Bedeutung hat. Diese Einschätzung ist wohl auch von Person zu Person verschieden. Tatsache ist, dass eine vollkommen wörtliche Wiedergabe aufgrund der technischen Möglichkeiten des Respeakings weitestgehend unmöglich scheint, weswegen genaues Paraphrasieren zweifelsohne eine wichtige Fähigkeit für RespeakerInnen ist, um Inhalte so genau wie möglich wiederzugeben.

3.4 Die Zielgruppe der Gehörlosen und Schwerhörigen

Um Live-Untertitelung zu verstehen, muss man auch verstehen, an wen sich diese hauptsächlich richtet. Wichtigste Zielgruppe und Grund für den im Moment stattfindenden Ausbau der Live-Untertitelung sind die Gruppe der Gehörlosen und Schwerhörigen (De Castro et al. 2015:51). Bei dieser Gruppe handelt es sich um eine diverse Gruppe mit unterschiedlichen Bedürfnissen, die von Faktoren wie zum Beispiel dem Alter zum Zeitpunkt des Hörverlustes, ob eine Gebärdensprache oder Lautsprache die Muttersprache ist, dem Stadium des Spracherwerbs vor dem Verlust des Gehörs und vielem weiteren abhängt (Eugeni 2008:40).

Allerdings ist dennoch wichtig zu erwähnen, dass gehörlose und schwerhörige Personen nicht die Einzigen sind, die Live-Untertitelung nutzen. Diese wird auch beispielsweise von Nicht-Muttersprachlern zum Spracherwerb genutzt oder dient zum besseren Verständnis von Sendungen in lauten Umgebungen (De Castro et al. 2015). Auch diese anderen Gruppen sind zahlenmäßig nicht unbedeutend: Einer Untersuchung des Ofcom zufolge haben von den 7,5 Millionen Menschen, die intralinguale Untertitel nutzen, 6 Millionen keine Hörbehinderung. Ähnliche Zahlen scheinen auch in anderen Ländern zutreffend zu sein (Waes & Remael 2013:3). Die Diversität dieser Zuhörerschaft sollte daher nicht vergessen werden, obgleich sich dieses Kapitel mit gehörlosen und schwerhörigen NutzerInnen befasst.

Wie bereits erwähnt, ziehen die meisten Gehörlosenorganisationen die wörtlichen Untertitel klar den editierten Untertiteln vor. Betrachtet man den Zweck der SDH, nämlich die Inklusion von Menschen mit Hörbehinderungen in den Medien, dann wird klar, wieso Editierung und Zusammenfassung in den Untertiteln als Exklusion und Zensur aufgefasst werden kann. Es gibt aber auch den Vorschlag, beiden Seiten gerecht zu werden, indem zwei Versionen der Untertitel gesendet werden. Eine editierte, vereinfachte Version für jene, die zum Beispiel aufgrund einer geringeren Lesegeschwindigkeit diese bevorzugen, und eine möglichst wortgetreue Version. Wirtschaftlich betrachtet ist das Senden zweier verschiedener Untertitelungen in den meisten Fällen allerdings kaum möglich, weswegen in der Praxis wohl ein anderer Kompromiss gefunden werden muss (Orero & Matamala 2010:12). In der wissenschaftlichen Diskussion hingegen wird unterstrichen, dass das Problem der Inklusion von der praktischen Anwendung in der Untertitelung getrennt betrachtet werden soll. Die wörtliche Untertitelung würde kein besseres Verständnis bedeuten, sondern kann im Gegenteil durch zu hohe Geschwindigkeit das Verständnis verschlechtern. Diese Annahme wird auch von Studien untermauert (Romero-Fresco 2009:111f).

3.4.1 Rezeptionsstudien

Bei Studien zu diesem Thema handelt es sich um sogenannte Rezeptionsstudien, mit denen – wie der Name bereits sagt – die Rezeption der Untertitel durch gehörlose und schwerhörige Personen normalerweise anhand von Experimenten, bewertet werden soll (Romero-Fresco 2018).

Die ersten dieser Studien gab es bereits mit Einführung der SDH in den 70er-Jahren. Diese Studien wurden über einen langen Zeitraum von der AVT-Forschung weitestgehend ignoriert. Sie befassten sich außerdem mit einem sehr eingeschränkten Themengebiet, nämlich

damit, wie Information am besten an gehörlose SchülerInnen vermittelt werden kann. Hierzu wurden verschiedene Modalitäten getestet und miteinander verglichen, darunter auch Untertitelung. Das wohl interessanteste Ergebnis dieser Untersuchung war die Erkenntnis, dass ein normales Redegeschwindigkeitstempo für die TeilnehmerInnen als Untertitelungstempo häufig zu schnell war. Daraus ergaben sich die ersten Berechnungen zur optimalen Untertitelungsgeschwindigkeit. Trotz dieser ersten Pionierstudien wurden die ersten großen Studien in Bezug auf die Rezeption von Fernseh-Untertiteln in den 80er-Jahren durchgeführt. Diese Studien ergaben zum einen, dass sich die untersuchten Personen mehr Untertitelung für Fernsehsendungen wünschten, und zum anderen unterschiedliche Ergebnisse zur möglichen Lesegeschwindigkeit, die stark vom Hintergrund der TeilnehmerInnen abhängig war. (Romero Fresco 2018:201ff).

Um die Jahrhundertwende entwickelte sich eine neue Art, diese Studien durchzuführen, nämlich die Eyetracking-Methode, die sich dazu eignete, genaue Aussagen über das tatsächliche Leseverhalten während des Konsums von Untertiteln zu erhalten. Frühe Ergebnisse besagten, dass die meisten Personen zunächst den Blick in die Mitte des Bildschirms richten, dann auf die Untertitel sehen, sobald diese angezeigt werden, um anschließend den Blick zurück zum Bild zu bewegen. Ein wichtiger Faktor für ein gleichwertiges Fernseherlebnis ist, wie viel Zeit auf das Lesen der Untertitel verwendet wird und wie viel Aufmerksamkeit im Ausgleich dazu für das Ansehen des Fernsehbildes verfügbar bleibt. Während die Ergebnisse diesbezüglich in den verschiedenen Studien variierten, konnte festgestellt werden, dass generell mit steigender Geschwindigkeit der Untertitel auch die Zeit, die auf die Untertitel aufgewendet werden muss, anstieg. Es wurden auch andere Aspekte, wie zum Beispiel die Wiedergabe der nicht-sprachlichen Informationen in der Untertitelung, erforscht (Romero Fresco 2018:207f).

Im europäischen Raum konnte sich Forschung zu diesem Thema erst später etablieren. Die ersten Untersuchungen zur Rezeption fanden bereits in den 90er-Jahren in Großbritannien statt, andere europäische Länder waren zu diesem Zeitpunkt aber kaum in der Forschung vertreten. Eine Wende hin zu einer breiteren wissenschaftlichen Untersuchung des Themas fand in Europa erst 2005 mit der Forschung von Joselia Neves zur Rezeption in Portugal statt. Weitere Arbeiten folgten und die Thematik begann auch in den europäischen Ländern präsenter zu werden (Romero-Fresco 208ff).

Eine der großen Rezeptionsstudien unter Gehörlosen und Schwerhörigen, die in den letzten Jahren durchgeführt wurde, ist in dem Buch „*The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*“, herausgegeben von Pablo Romero-Fresco, beschrieben. Es handelte sich um eine groß angelegte Studie eines Projektes der Europäischen Kommission

namens DTV4ALL. Ziel dieses Projektes war es, die Untertitelung aus Sicht der NutzerInnen zu analysieren, sowohl in Bezug auf deren Verständnis des Gezeigten als auch der Rezeption. Daraus sollten dann in weiterer Folge Richtlinien für die Untertitelung entstehen. Hierbei wurden gehörlose und schwerhörige Personen in verschiedenen europäischen Ländern unter anderem mittels Fragebögen befragt.

Untersuchte Themengebiete waren die demografischen Informationen und die Fernsehgewohnheiten der Befragten, die Qualität, die Wörtlichkeit und die Geschwindigkeit der Untertitel, sowie die formalen Kriterien, die man an die Untertitel stellte. Hierbei stellte sich heraus, dass in den meisten Fällen eine Geschwindigkeit von 150 wpm zu einer gleichmäßigen Verteilung der Aufmerksamkeit zwischen Bild und Untertiteln führte, während bei der höheren Geschwindigkeit von 180 wpm mehr Zeit für die Untertitel benötigt wurde. Auch lag das Verständnis der hörbehinderten TeilnehmerInnen durchschnittlich 15 Prozent niedriger als bei Hörenden, was auf Leseschwierigkeiten hindeuten kann. Zugleich fiel aber auch auf, dass trotz der kürzeren Zeit, die Gehörlose und Schwerhörige für das Betrachten der Bilder aufwenden konnten, ihr visuelles Verständnis dennoch sehr hoch blieb. Weiteres Ergebnis der Studien war erneut, dass die Situation der Zielgruppen in den einzelnen Ländern sehr heterogen ist. Kulturelle und soziale Faktoren spielen eine große Rolle, weswegen eine einheitliche Regelung für die Untertitelung in ganz Europa zumindest zum momentanen Zeitpunkt eher nicht möglich sein wird (Romero-Fresco 2015).

Es gibt nach wie vor nur vereinzelte dieser Rezeptionsstudien, die sich rein mit der Live-Untertitelung befassen. Diese beschränken sich meistens auf Nutzerbefragungen. Erste auf diese Art durchgeführte Studien von Eugeni und Romero-Fresco unter italienischen beziehungsweise britischen NutzerInnen konnten bereits wertvolle Ergebnisse bezüglich der präferierten Geschwindigkeit, Wörtlichkeit und Anzeigemodalitäten liefern. Um genauere und aussagekräftigere Ergebnisse zu liefern, die beispielsweise den Fernsehsendern bei der Erstellung von an die Bedürfnisse der NutzerInnen angepasster Untertitelung hilfreich sein könnten, bedarf es noch weiterer Forschung auf diesem Gebiet (Romero-Fresco 2018:215).

4 Respeaking in Japan

In diesem Kapitel werden allgemeine Informationen zur Gehörlosigkeit in Japan und der gesetzlichen Situation gegeben. Besonders eingegangen wird hier auf die Richtlinien des japanischen Innenministeriums für die Live-Untertitelung. Anschließend befasst sich das Kapitel mit den verfügbaren Artikeln zum Thema der Live-Untertitelung bei NHK und fasst zusammen, wie die Geschichte und die heutige Verwendung der Untertitelung und insbesondere der Respeaking-Methode aussieht.

4.1 Gehörlose in Japan

Laut einer Umfrage des japanischen Ministeriums für Gesundheit, Wohlfahrt und Arbeit (厚生労働省) gab es im Jahr 2016 etwa 340 000 Menschen mit einer Hörbehinderung in Japan. Hinzu kommt, dass Japan durch seine überalterte Gesellschaft eine Zunahme an altersbedingter Schwerhörigkeit zu verzeichnen hat (Ministerium für Gesundheit, Wohlfahrt und Arbeit 2018). Den Schutz der Rechte von Menschen mit Behinderungen gewährleisten in Japan das Grundgesetz über Menschen mit Behinderung [jap: 障害者基本法], das im Jahr 1970 erlassen und 2011 zuletzt reformiert wurde, und das Gesetz zur Eliminierung von Diskriminierung gegen Personen mit Behinderungen [jap: 障害を理由とする差別の解消の推進に関する法], das im Jahr 2013 erlassen wurde. Dieses Gesetz sichert Menschen mit Behinderung dieselben Rechte zu, wie sie Personen ohne Behinderung haben, und verbietet es staatlichen und privaten Organisationen, auf Basis von Behinderungen zu diskriminieren.

Die nachträgliche Ergänzung des Grundgesetzes betrifft das Entfernen von Barrieren für behinderte Personen in der Gesellschaft, sprich die Teilnahme derselben an dieser. So heißt es im Grundgesetz:

When a person with a disability currently requires the removal of a social barrier and if the burdens associated with said implementation is not excessive, necessary and reasonable accommodation must be given to implementing the removal of the social barrier so as not to be in violation of the provisions of the preceding paragraph [Grundrechte von Menschen mit Behinderung, Anm.] by denial to do so.

(Basic Act for Persons with Disabilities 2011/90 §4 Abs. 2 [offizielle engl. Übersetzung]):

Das Gesetz zur Eliminierung von Diskriminierung gegen Personen mit Behinderungen kolportiert in ähnlicher Formulierung, dass Barrieren für Personen mit Behinderungen in allen Bereichen eliminiert werden sollten. (Act for Eliminating Discrimination against Persons with Disabilities 2013/65) Diese Anweisungen sind zwar nicht besonders konkret formuliert,

entsprechen aber in etwa auch den Forderungen, wie sie in der UN-Behindertenkonvention angegeben sind. Diese UN-Konvention, die in Japan 2014 ratifiziert wurde, bildete auch eine Basis für die Gesetzes-Änderungen in Japan (Innenministerium 2017:2). Einen Teil dieser Inklusion aller Personen in die Gesellschaft macht natürlich auch die Partizipation in den Medien aus. Daher wurden auch Anstrengungen unternommen, die Integration in diesen Bereichen zu verbessern.

Einer Umfrage des Japanischen Gehörlosenverband zufolge haben 88 Prozent aller hörbehinderten Personen einen Fernseher zuhause. (Ministerium für Gesundheit, Wohlfahrt und Arbeit 2018). Aufgrund der Bestrebungen für mehr Barrierefreiheit ergriff das japanische Innenministerium Initiativen, um die Barrierefreiheit für alle auch im Rundfunk zu gewährleisten, so auch für jene Personen, die nicht oder nur schlecht hören können. Auch speziell zur Barrierefreiheit in den Medien gibt es in Japan gesetzliche Grundlagen. Im Jahr 1997 wurde das japanische Rundfunkgesetz [jap: 放送法] erlassen. Laut §4 dieses Gesetzes müssen so viele Sendungen wie möglich durch Untertitelung oder andere Maßnahmen gehörlosen oder hörbehinderten Menschen zugänglich gemacht werden (Broadcasting Act 2014/96, Ito 2013:950). Im selben Jahr brachte das damalige Post- und Telekommunikationsministerium eine Richtlinie heraus, laut welcher innerhalb der nächsten 10 Jahre alle vorab aufgezeichneten Sendungen beim öffentlich-rechtlichen Fernsehsender NHK untertitelt sein sollten. Dieses Ziel erreichte NHK im Jahr 2006. Als weiterer Schritt nach diesem Ausbau folgte 2008 die erste Richtlinie des Innenministeriums (wie weiter unten besprochen), nach welcher bis zum Jahr 2017 alle Sendungen inklusive Live-Sendungen, die zwischen 7 und 24 Uhr ausgestrahlt werden, untertitelt sein sollten (Innenministerium 2008, 2017).

Einer großen Mobilisierungswelle in Bezug auf die Barrierefreiheit des Fernsehens lag die Erdbebenkatastrophe in Ostjapan im Jahr 2011 zugrunde. Obwohl die Medien auch zuvor schon angehalten waren, ihre Übertragungen für alle BürgerInnen zugänglich zu machen, kamen im Nachhinein der Katastrophe erschreckende Zahlen ans Licht. Während in den besonders stark betroffenen Gebieten in der Allgemeinbevölkerung die Sterblichkeitsrate bei etwa 1,03% lag, war der Anteil bei Personen mit einer Hörbehinderung mit 2% doppelt so hoch. Da ein großer Anteil dieser Personen physisch in der Lage gewesen wäre, sich an einen sicheren Ort zu begeben, liegt die Vermutung nahe, dass zumindest in einige Fällen mangelnde Informationsvermittlung ein Grund für diese erschreckenden Zahlen sein könnte (Ito 2013: 949).

Das japanische Innenministerium gab 2018 nach einer Konferenz, an welcher auch Vertreter der Interessengemeinschaften gehörloser Menschen teilgenommen hatten, eine neue

Richtlinie heraus, die den Weg in die Zukunft weisen soll. Laut ihren Erhebungen konnte die hundertprozentige Untertitelungsrate nicht erreicht werden, auch wenn die Rate je nach Fernsehsender relativ hoch war (Innenministerium 2017).

So lag im Jahr 2016 die Untertitelungsrate der einzelnen Fernsehsender bei Werten zwischen 97 und 99,5 Prozent. Es ist aber zu beachten, dass nur Sendungen untersucht wurden, die innerhalb der von den Richtlinien vorgegebenen Zeiten ausgestrahlt werden. Während die Untertitelungsrate bei großen Fernsehsendern also relativ hoch ist, ist sie hingegen bei den Lokalsendern nach wie vor niedrig. Dies wird als Problem angesprochen, das schwierig zu lösen ist, da die Lokalsender häufig über ein eher niedriges Budget verfügen und daher teure Untertitelungsmethoden, besonders im Bereich der Live-Untertitelung, für sie unerschwinglich sind. Ein weiterer in der Richtlinie erwähnter Vorschlag war eine Untertitelung auch für Werbeunterbrechungen, wozu bereits einige Experimente durchgeführt wurden. Ebenso wurde die Wichtigkeit der staatlichen Fördergelder für die Untertitelung angesprochen, da diese es ermöglichen, auch zum Beispiel bei kleineren Sendern mehr untertitelte Sendungen anzubieten (Innenministerium 2017).

4.2 Nutzung von Spracherkennung zur Live-Untertitelung bei NHK

Es existieren mehrere in japanischer Sprache verfasste Artikel zur Entstehungsgeschichte und Anwendung der Live-Untertitelung in Japan. Diese beschäftigen sich allerdings fast ausschließlich mit der Verwendung bei NHK und werden normalerweise auch von dessen internem Forschungszentrum herausgegeben. Es ist daher schwieriger, Aussagen zu den Privatsendern zu machen, weswegen ich auch beschlossen habe, mich für diese Arbeit auf NHK zu konzentrieren. Warum dieses Material in der Arbeit ausführlich besprochen wird, hat im Wesentlichen zwei Gründe: zum einen bieten diese Artikel eine gute Grundlage, um die Grundzüge der Live-Untertitelung im öffentlich-rechtlichen Fernsehen in Japan zu verstehen, und um in der darauffolgenden empirischen Forschung auf detaillierte Aspekte und spezifische Fragen einzugehen. Zweitens sind die darin enthaltenen Informationen zum großen Teil nur in japanischer Sprache verfügbar, weswegen sie den LeserInnen dieser Arbeit durch die detaillierte Wiedergabe zugänglich gemacht werden sollen.

4.2.1 Anfänge der Spracherkennung

Beim japanischen Fernsehsender NHK wird seit 1985 Untertitelung im Fernsehen verwendet, die vor allem für gehörlose und schwerhörige Personen als Zielgruppe intendiert ist. Diese sollte zunächst vordergründig dazu dienen, diesen Personen den Zugang zu wichtigen Informationen, insbesondere in Form von Nachrichten- und Informationssendungen, zu ermöglichen. Später wurde das Angebot über die folgenden Jahre hinweg auch auf andere Formate, wie Serien und Dokumentationen, ausgeweitet. Aufgrund der fehlenden technischen Möglichkeiten konnte jedoch zum damaligen Zeitpunkt noch keine Untertitelung bei Live-Sendungen stattfinden (Homma 2010:25).

Bereits im Jahr 1996 wurde mit Forschung zum Einsatz von Spracherkennung für Live-Untertitel begonnen. Zunächst wurden Aufnahmen von Fernsehsprechern und im Voraus erstellte Drehbücher von Sendungen gesammelt, um aus ihnen eine Datenbank zu erstellen. Diese konnte die Genauigkeit der Spracherkennung so weit verbessern, dass erste Experimente in der Untertitelung von Live-Sendungen möglich waren. Im März 2000 wurden dann zum ersten Mal Live-Untertitel, die sich auf die vom Nachrichtensprecher vorgelesenen Teile der Nachrichtensendung „News 7“ beschränkten, ausgestrahlt.

Es ist allerdings zu erwähnen, dass für diese Untertitel kein Respeaking, also keine Wiederholung des Gesprochenen eingesetzt wurde, sondern die Stimme des Nachrichtensprechers direkt erkannt wurde. Es gab hierbei zwei separate Spracherkennungssysteme, eines für weibliche und eines für männliche Stimmen, jeweils mit eigenen Editoren. Diese strikte Trennung hängt wohl vor allem mit dem im Japanischen sehr unterschiedlichen Sprachgebrauch von Männern und Frauen zusammen. Das System verfügte auch über ein Backup-System als Redundanz und war allgemein in seiner Konstruktion eher kompliziert. Bei der unten Abbildung 4 handelt es sich daher um eine Verkürzung, in der diese Aspekte ausgelassen und nur das Spracherkennungssystem als solches dargestellt wurde.

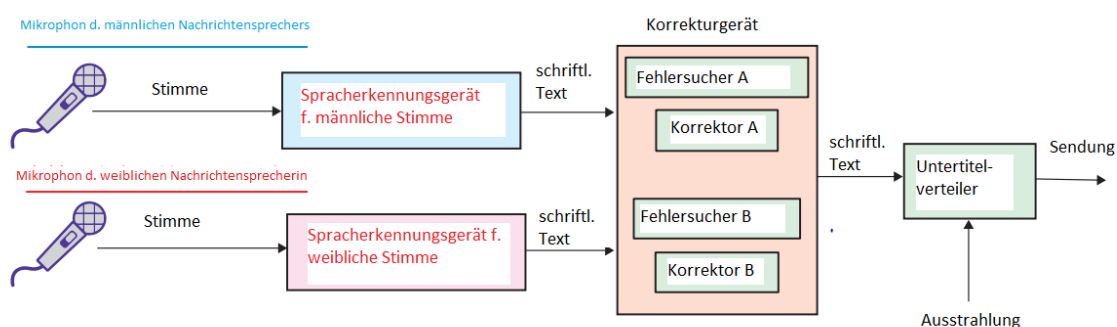


Abb. 4: Untertitelungssystem (Original Homma 2010:26, Eigene Übersetzung)

Es waren pro System (männliche beziehungsweise weibliche Stimmerkennung) also jeweils zwei Personen für die Korrektur der Untertitelung beschäftigt, eine Person als „Fehlersuchende“ und eine weitere für die tatsächliche Korrektur. Der Ablauf war also so strukturiert, dass die „fehlersuchenden Personen“ die inkorrekten Stellen auf einem Touchpad markierten und so die KorrektorInnen aufmerksam machten, die wiederum die Fehler über ein Tastatur ausbesserten. Da dieses System der direkten Erkennung auf diese Art natürlich einen großen Aufwand mit sich brachte, insbesondere da insgesamt vier MitarbeiterInnen für die Korrektur beschäftigt waren, wurde es im Jahr 2006 wieder eingestellt und durch die Nutzung einer Schnellschreibtastatur ersetzt (Homma 2010:26, Homma et al. 2009:331).

Wie bereits erwähnt, funktionierte die Methode der direkten Spracherkennung auch nur bei bestimmten Sendungen beziehungsweise Teilen dieser Sendungen. Bei Nachrichtensendungen konnten zum Beispiel der abgelesene Text des Nachrichtensprechers, sowie zum Teil auch die Aussagen von Reportern vor Ort mit einer zufriedenstellenden Exaktheit wiedergegeben werden. Bei den häufig von Hintergrundgeräuschen unterbrochenen und in Umgangssprache gehaltenen Aussagen beispielsweise von auf der Straße interviewten Laien brachte die Methode allerdings keine verwertbaren Resultate mehr hervor (Homma 2010: 27). Es war jedoch das Ziel des Senders, eine durchführbare und möglichst kostengünstige Methode zu finden, die Live-Untertitelung auch auf andere Programme, wie zum Beispiel Sportsendungen oder Musiksendungen, auszuweiten. Daher wurden gegen Ende des Jahres 2001 die ersten Sendungen mithilfe von Respeaking untertitelt. Beispiele sind der Fußball World Cup und professionelles Baseball (Homma 2010:26).

4.2.2 Respeaking-Methode bei NHK

Um also eine kostengünstige und einsatzstarke Alternative zur Verwendung von Steno-Keyboards zu schaffen, wurde die Spracherkennung durch Respeaking als Methode in Betracht gezogen und auch getestet. Als Gründe für die Anwendung der Respeaking-Methode werden folgende Vorteile aufgelistet: erstens sei es mit dieser Methode möglich, auch Sendungen mit lauten Hintergrundgeräuschen und mehreren Sprechern zu untertiteln. Durch die Möglichkeit der Nutzung eines an den Respeaker angepassten Sprachmodells könne die Genauigkeit der Erkennungsrate gesteigert werden. Darüber hinaus können Inhalte wie Klatschen oder Ausrufe wiedergegeben, sowie Inhalte, die durch das Fernsehbild ohnehin offensichtlich sind, verkürzt werden. Außerdem können Fehler in der Spracherkennung durch Umformulierungen verhindert werden. Als Nachteile der Methode werden hingegen die größere Verzögerung, die fehlende

Garantie einer vollständigen Wiedergabe durch den/die RespeakerIn und die Kosten für das notwendige Personal genannt (Homma 2010:26).

Im Jahr 2006 lag die Untertitelungsrate abzüglich der immer noch schwierigen Live-Sendungen bei NHK bereits bei 100 Prozent. Bei den fünf größten Privatsendern lag sie derweil im Durchschnitt immerhin bei 77.8 Prozent. Diese sehr guten Werte sinken allerdings drastisch, sobald man die Live-Sendungen miteinbezieht. Rechnet man diese mit ein, so lag NHK bei 43 Prozent und alle anderen Sender bei etwa 32 Prozent (Innenministerium 2008).

Ein besonderer Fokus bei der Untertitelung wurde auf Nachrichtensendungen und Übertragungen im Katastrophenfall gelegt (Homma et al. 2009:332). Für diese Zwecke wurden vor allem die bereits erwähnten Schnellschreibtastaturen eingesetzt. Problematisch ist hierbei aber neben den Kosten auch die Verfügbarkeit. Aufgrund der langen Dauer der Ausbildung gibt es nicht viele Personen, die die erforderliche Technik beherrschen. Gerade in Krisenzeiten werden 24 Stunden am Tag wichtige Informationen gesendet, für deren Transkription es dann einfach nicht genug qualifizierte Kräfte gibt (Innenministerium 2017).

In einem 2007 veröffentlichten Artikel schlägt Shinichi Homma aus dem NHK-eigenen Forschungszentrum eine neue Methode vor, um mit den Problemen der direkten Spracherkennung umzugehen, Respeaking miteinzubinden und die Kosten gering zu halten. Diese Methode bezeichnet er als die „Hybridmethode“ und sie ist, wie der Name bereits anzeigt, ein Hybrid aus der direkten Spracherkennung und dem Respeaking. Bei dieser Methode wird der Audio-Input des/der NachrichtensprecherIn direkt erkannt und in Text umgewandelt. Sobald jedoch zum Beispiel eine Live-Schaltung oder Ähnliches stattfindet, wo Probleme in der Erkennung durch Hintergrundgeräuschen oder mehrere sich überlappende Sprecher entstehen können, wird die Spracherkennung auf eine/n sich im Studio befindliche/n RespeakerIn umgeleitet. (Homma et al. 2007:1, 2009:332). Die RespeakerInnen haben hierbei die Aufgabe, den Output der Untertitelung zu überwachen und bei für die direkte Erkennung problematischen Inhalten auf die Erkennung durch Respeaking umzuschalten. Damit bei der Rückschaltung auf die automatische Erkennung kein Audio-Input verloren geht, ist im System ein Puffer von einer Sekunde eingebaut.

Die Vorteile dieses Systems benennt Homma wie folgt: erstens könne so die Anwendung auf andere Sendungsarten ausgeweitet werden und die Nachrichtenuntertitel können vollständig durch Spracherkennung erstellt werden, ohne dass beispielsweise auf ein Schnellschreib-Tastatur zurückgegriffen werden müsse. Durch eine Verbesserung des Korrektursystems seien jetzt statt vier nur noch ein bis zwei Personen für die Korrektur notwendig. Darüber hinaus sei ihre Ausbildung, da Touchscreens oder normale Tastaturen verwendet werden, wesentlich

weniger aufwändig als die der bisher eingesetzten StenotypistInnen. (Homma et al. 2007, Homma 2010:27).

Zeitgleich mit der Einführung der hybriden Methode, wurde auch die Spracherkennung als solche verbessert. Das System konnte nun eigenständig zwischen männlichen und weiblichen Stimmen unterscheiden. Dies hat den Vorteil, dass nicht mehr zuvor das Geschlecht des Sprechers eingegeben werden muss, und auch keine zwei getrennten Systeme mehr nötig sind. Das Hybridsystem besteht somit aus weniger Komponenten als sein Vorgänger. Ebenso lernt das Sprachmodell mit Hilfe eines verdeckten Markovmodells aus diesen Texten die Wahrscheinlichkeiten des Aufeinanderfolgens von Worten oder Satzteilen, was eine automatische Textvervollständigung ermöglicht. Dies erlaubt es dem System, Aussagen bereits vor Beendigung des Satzes vorherzusehen und mit hoher Wahrscheinlichkeit korrekt zu verschriftlichen. (Homma et al. 2009:332). Ergänzt wird das System durch ein Phonetik-Wörterbuch, in dem die Aussprache wichtiger Vokabel festgehalten ist. (Homma 2010:28) Das Korrektursystem verfügte später auch über eine Merkfunktion, in der vorherige Korrekturen gespeichert und dann wieder als Vorschläge angezeigt werden (Satou 2018:164).

Zusätzlich zur Spracherkennung wurde auch das Korrektursystem verbessert. Während für das alte System eine fehlersuchende und eine korrigierende Person notwendig war, konnte beides nun von nur einer Person übernommen werden, die zunächst die falschen Outputs auswählt und sie dann auch selbst korrigiert. Die Anzahl der KorrektorInnen konnte je nach Schwierigkeitsgrad der Sendung angepasst verringert werden, wobei die Minimalanzahl nunmehr von zwei auf eine Person reduziert werden konnte. Wenn zwei oder mehr EditorInnen eingesetzt werden, erscheint der zu korrigierende Text in unterschiedlichen Farben, je nachdem wessen Aufgabe die Korrektur der besagten Stelle ist. Der/die EditorIn hat die Möglichkeit, sich eventuell unklare Stellen im Text erneut anzuhören. Die dadurch verlorene Zeit wird dann durch eine beschleunigte Wiedergabe des Folgetextes wieder ausgeglichen. Die EditorInnen bekommen außerdem im Fall von Homonymen die möglichen Bedeutungen als Liste vorgeschlagen und können schnell mithilfe eines Touchpads aus ihnen wählen, ohne die Tastatur nutzen zu müssen (Homma et al. 2009:334, Homma 2010:27f, Satou 2015:302f).

In einem von Homma durchgeführten Experiment, bei welchem eine reale Nachrichtensendung einmal mit dem alten und einmal mit dem neuen System untertitelt wurde, ergab sich, dass das neue System eine große Verbesserung in der Genauigkeit der Untertitelung erreichte. Nicht nur konnte nun der gesamte Inhalt der Sendung nur mit Spracherkennung untertitelt werden und nicht mehr nur der Text, der von den NachrichtensprecherInnen abgelesen wird; darüber hinaus sank die WER von 8,2 Prozent auf 3,8 Prozent. Die

Verzögerung der Untertitel lag durchschnittlich bei 5 Sekunden und wies damit keinen signifikanten Unterschied mehr zu jener der Schnellschreibtastatur-Methode auf. Im Fall von Interviews mit mehreren Personen oder Sendungsteilen mit lauten Hintergrundgeräuschen sank die Genauigkeitsrate allerdings beachtlich ab, und eine Korrektur war teilweise weiterhin nur schwer möglich. (Homma et al. 2009:335)

Generell war eine Kombination aus der Respeaking-Methode und der Direkterkennung am effektivsten, da die Respeaking-Methode bei einem zu langen Einsatz fehleranfälliger wurde, die direkte Methode dafür allerdings bei gewissen Programmteilen völlig versagte. Die Hybridmethode kann daher als gute Alternative zur Steno-Tastatur-Methode betrachtet werden (Homma et al. 2009:336f). Diese Möglichkeit galt aber zu diesem Zeitpunkt nur für Sendungen, in welchen sich durch direkte Methode übersetzbare Sendungsteile mit solchen, die Respeaking notwendig machten, abwechselten. Bei Sendungen, die keine Sendungsabschnitte aufwiesen, bei denen zur direkten Erkennung gewechselt werden kann, kam es durch einen zu langen Einsatz von Respeaking als einziger Untertitelungsmethode irgendwann zu so starken Verzögerungen, dass eine vollständige Untertitelung unmöglich wurde. Dadurch blieben immer noch etwa 30 Prozent der im Fernsehen ausgestrahlten Sendungen nicht untertitelbar (Homma 2010: 29).

Einen weiteren Aufschwung erlebte die Untertitelung im japanischen Fernsehen durch die im Jahr 2012 vom japanischen Innen- und Kommunikationsministerium herausgegeben Ergänzung zur Richtlinie zur Barrierefreiheit im Rundfunk aus dem Jahr 2008. In dieser Ergänzung, die eine Antwort auf die Erdbebenkatastrophe im März 2011 darstellt, wurde noch einmal gesondert darauf hingewiesen, wie im Katastrophenfall die Barrierefreiheit der Sendungen gewährleistet werden sollte. NHK wurde angehalten, im Falle von Katastrophen großen Ausmaßes so weit als möglich für alle Notfall-Übertragungen und alle regulären Nachrichtensendungen eine komplette Untertitelung bereitzustellen. Auch die Sendungen mit Gebärdensprachdolmetschung sollten erweitert werden (Innenministerium 2012). Diese Richtlinie diente als die Grundlage der Untertitelung bei NHK, und ihre Ergänzung als eine weiterer Impetus, die Untertitelung zu verbessern. (Ito 2013: 949).

Im Jahr 2015 wurden laut eines Artikels von Shoei Satou bei NHK vier verschiedene Methoden zur Untertitelung eingesetzt, die er wie folgt auflistet:

1. Die Nutzung einer herkömmlichen Tastatur, wobei mehrere UntertitlerInnen in einem Relais zusammenarbeiten. Obwohl diese Methode den Vorteil hat, ohne eine spezialisierte Ausbildung der UntertitlerInnen auszukommen, können aufgrund der

langsamen Geschwindigkeit nur Sendungen mit wenig gesprochenen Teilen, wie etwa Musiksendungen, auf diese Weise Untertitelt werden.

2. Der Einsatz der oben beschriebenen Steno-Tastatur. Diese Methode lässt sich auf eine Vielzahl von Sendungen mit großer Genauigkeit anwenden, ist aber aufgrund der langen Ausbildung sehr teuer, und es ist schwierig, qualifizierte Kräfte zu finden.
3. Die Respeaking-Methode, inklusive des Einsatzes von KorrektorInnen. Diese wird beispielsweise bei Sportsendungen wie Sumowettkämpfen eingesetzt und bei Talkshows.
4. Die direkte Spracherkennung bei jenen Texten, wo dabei eine hohe Genauigkeit zu erwarten ist. Diese wird normalerweise im Kontext der sogenannten Hybridmethode verwendet, bei der jene schwierigeren Sendungsteile, bei denen beispielsweise Vor-Ort-Interviews durchgeführt werden, von RespeakerInnen übernommen werden. Der Einsatz der Respeaking-Methode bleibt bei diesen Sendungen normalerweise sehr gering, da viele Vor-Ort-Segmente im Voraus aufgenommen werden und daher auch im Voraus Untertitelt werden können. Hinzu kommt, dass solche Segmente im japanischen Fernsehen tendenziell ohnehin mit *open captions* Untertitelt sind.

Bei der Respeaking- und der Hybrid-Methode hängt die Genauigkeit stark davon ab, wie gut das System an die jeweilige Sendung angepasst ist, insbesondere in Bezug auf das eingelesene Vokabular. Auch die Geschwindigkeit des gesprochenen Textes spielt in Zusammenhang mit der Genauigkeitsrate eine Rolle. Wird die Genauigkeit zu niedrig und ist die Geschwindigkeit zugleich zu schnell, können die KorrektorInnen nach wie vor nicht mehr genügend Fehler ausbessern, um eine qualitativ hochwertige Untertitelung zu gewährleisten. (Satou 2015:300)

Zum Zeitpunkt der von Satou durchgeführten Studie wurden etwa die Hälfte aller Live-Sendungen mit einer Form der Spracherkennung Untertitelt. Bedingt durch die häufig emotionale Ausdrucksweise der Sprecher und die vielen Hintergrundgeräusche kommt für den Einsatz bei Sportsendungen die direkte Spracherkennung nicht in Frage. Beim Einsatz der Respeaking-Methode für Sportsendungen lag die Verzögerung bei durchschnittlich 5-10 Sekunden. Diese Verzögerung muss natürlich miteinberechnet werden, da gerade bei einem sportlichen Wettkampf das auf dem Bildschirm Gezeigte mit den Untertiteln zusammenpassen muss, um sinnhaft zu sein. Hier wenden die UntertitlerInnen teilweise Verkürzungen des Textes an, um diese zeitlich wieder an das Geschehen anzupassen. Genauso müssen die RespeakerInnen natürlich auch im Falle von mehreren Personen, die zugleich sprechen, wie das zum Beispiel im Falle von Varietyshows und Talkshows häufiger der Fall ist, die verschiedenen

Sprecher zum besseren Verständnis kennzeichnen. Der/die RespeakerIn wird laut Satou bei dieser Art der Untertitelung etwa alle 10 Minuten ausgewechselt, und es kümmern sich ein bis zwei KorrektorInnen um das Ausbessern des Textes. Die RespeakerInnen verfügen dabei individuell über ein auf ihre Aussprache trainiertes Sprachprofil, das innerhalb von Sekunden umgeschaltet werden kann. Bei der direkten Erkennung kann natürlich kein speziell auf den Sprecher zugeschnittenes Modell zum Einsatz kommen. Wie bereits oben erwähnt besitzt das System allerdings die Fähigkeit, selbst zwischen einem weiblichen und männlichen Sprecher zu unterscheiden (Satou 2015:302, 2018:164f).

Abgesehen von den individuellen Sprecherprofilen verfügt das System auch über eigene Sprachmodelle und Vokabeldatenbanken für die verschiedenen Sendungsgenres. Das System kann auch selbstständig auf zuvor eingelesene Nachrichtenartikel oder Drehbücher zu den besprochenen Themen zugreifen und aus ihnen zum Beispiel Orts- oder Personennamen abgleichen. Eine große Schwierigkeit stellen hierbei aber manche Informationssendungen und Talkshows dar. Nicht nur haben diese einen sehr breiten Umfang an möglichen Themen, gerade bei letzterem Genre gibt es meistens auch im Voraus kein detailliertes Drehbuch, sondern nur einen ungefähren Ablauf. Dadurch ist es schwieriger, gute Trainingstexte für die Spracherkennungssoftware zu finden.

Eine Möglichkeit, diesen Mangel auszugleichen, liegt aber in den häufig vor den Sendungen stattfindenden Proben. Der für die Sendung Verantwortliche kann hierbei eine Art Drehbuch zusammenstellen; darüber hinaus kann die Probe selbst bereits durch Respeaking wiedergegeben werden, wodurch mögliche Problembereiche augenfällig werden. Auch während der tatsächlichen Sendung kann der/die RespeakerIn den produzierten Output der Spracherkennung im Auge behalten und so eventuelle Fehler gleich durch eine Neuformulierung des jeweiligen Satzes ausgleichen, anstatt dass diese erst von den EditorInnen angepasst werden müssen (Satou 2015:302).

Auch die Vorbereitung der RespeakerInnen und EditorInnen wird in den Arbeiten erwähnt. Als Vorbereitung für die Sendung sehen sich die RespeakerInnen und EditorInnen die in Voraus veröffentlichten Drehbücher der Nachrichtensendungen an. Mithilfe dieser stellen Sie fest, welche Inhalte mittels Respeaking wiedergegeben werden müssen, und lesen neue oder schwierige Vokabel in das System ein. Die EditorInnen überprüfen darüber hinaus die

Funktionstüchtigkeit der Kana-Kanji Konversion bei problematischen Begriffen¹. (Homma et al. 2009:334).

Eine Fragestellung, mit der sich Satou in seinem Text beschäftigt, ist auch erneut jene der lokalen Nachrichtenstation. Während in den größeren Sendezentralen die oben besprochene Methode der hybriden Untertitelung Anwendung findet, stehen Lokalsender vor einem Kostenproblem bei der Etablierung. Nicht nur sind die Anschaffungskosten für das neue System sehr hoch, es fehlen auch qualifizierte Kräfte, um die Korrektur zu übernehmen. Daher wurde für diese Zwecke eine eigene Methode entwickelt, bei der das System weitestgehend automatisch arbeiten kann.

Da die meisten Sendungen, insbesondere Nachrichtensendungen, auf einem Drehbuch basieren, wird das Drehbuch zuvor ins System eingelesen. Sollten sich kurzfristige Änderungen kurz vor oder sogar während der Sendung ergeben, können diese noch von einem/r MitarbeiterIn korrigiert werden, wozu keine besondere Ausbildung vonnöten ist. Die Spracherkennungssoftware wandelt den gesprochenen Text direkt um, und dieser wird mit dem Drehbuch verglichen. Aufgabe des Systems ist es hierbei, die passende Stelle im Drehbuch zu finden und als Untertitel zu senden. Wenn der/die SprecherIn den Inhalt verändert, wird diese Änderung allerdings nicht wiedergegeben. Genauso wenig kann das System Sendungsteile wiedergeben, die nicht im Drehbuch stehen. Diese müssen entweder von den offenen Untertiteln wiedergegeben werden oder fehlen eben komplett. Daher kann diese Methode nicht immer eine vollständige Wiedergabe erreichen, noch ist sie für Talkshows und ähnliche Sendungen ohne genaues Drehbuch anwendbar. Die Vorteile sind die Einsparung von Kosten und die einfache Anwendung.

In einem 2018 erschienenen Artikel gibt Shoei Satou die aktuelle Situation der Untertitelung bei NHK wieder. Hierbei handelt es sich um die aktuellste wissenschaftliche Publikation zu dieser Thematik in Japan, auf die ich während meiner Recherche Zugriff nehmen konnte. Laut Satou (2018) sind zu diesem Zeitpunkt 97,4% aller Sendungen untertitelt. Hierbei geht aber nicht ganz klar hervor, ob es sich wirklich um das gesamte Spektrum an ausgestrahlten Sendungen handelt oder nur um jene, die zwischen 7:00 und 24:00 Uhr ausgestrahlt werden, da Satou sich auf die Richtlinie des Innenministeriums bezieht, die ja nur Zielsetzungen für diese Sendezeiten vorgibt. Jedenfalls lässt sich daraus schließen, dass die Anzahl an untertitelten

¹ Da es unmöglich ist, alle japanischen Schriftzeichen auf einer Tastatur dazustellen, werden auf herkömmlichen Tastaturen entweder Kana (Symbole der japanischen Silbenschrift) oder lateinische Buchstaben eingegeben, welche dann vom System in die passenden Schriftzeichen umgewandelt werden. In diesem Fall geht es darum, dass das System das korrekte Wort auswählt anstatt eines der (in vielen Fällen recht zahlreichen) Homonyme.

Sendungen im Vergleich zu anderen Ländern sehr hoch ist. Allerdings beziehen sich diese Zahlen natürlich nur auf NHK selbst und nicht auf Regional- oder Lokalsender.

Im Jahr 2018 wurden alle fünf in den vorherigen Absätzen beschriebenen Methoden verwendet. Das bedeutet, dass je nach Sendungstyp und Sender folgende Varianten zur Live-Untertitelung in folgenden Prozentsätzen angewendet wurden:

- 1) Die Relais-Methode, für die eine konventionelle Tastatur verwendet wird.
- 2) Die Steno-Tastatur-Methode. Diese macht gemeinsam mit der Verwendung einer normalen Tastatur etwa 50 Prozent der untertitelten Live-Sendungen aus.
- 3) Die reine Respeaking-Methode, die für etwa 40 Prozent der Sendungen verwendet wird.
- 4) Die Hybrid-Methode, welche nur 10 Prozent ausmacht.
- 5) Die Drehbuch-Vergleichsmethode, die aber nur in geringem Ausmaß und nur bei Regionalsendern angewendet wird.

Für die Untertitelung mit der Steno-Tastatur werden bei NHK Arbeitskräfte einer Fremdfirma beauftragt. Obwohl die Untertitelung mittels Respeaking weniger langwierige Ausbildung benötigt, wird auch hier das Finden von qualifizierten RespeakerInnen von Satou als Problem angeführt (Satou 2018:165).

Zusammengefasst wird die Respeakng-Methode in der Live-Untertitelung in Japan bereits seit den frühen 2000er-Jahren eingesetzt. Einen weiteren Impetus für die Entwicklung stellte die Erdbebenkatastrophe 2011 dar. Während am Anfang die direkte Spracherkennung beziehungsweise eine Steno-Methode verwendet wurde, fand Respeaking schon relativ früh Einzug in die Live-Untertitelung, da es, wie in vielen anderen Ländern auch, als leistungsstarke und wenig kostenintensive Methode angesehen wurde. RespeakerInnen arbeiten in Japan wegen der komplizierten Schriftumwandlung immer mit mindestens einem, meistens mehreren EditorInnen zusammen. Neben dem Respeaking sind weiterhin auch die direkte Spracherkennung, die Steno-Tastaturmethode und eine Methode, bei der ein Drehbuch als solches gesendet wird, in Verwendung.

5 Methodik

Auf die theoretische Bearbeitung des Themas Respeakings in Japan folgt nun der empirische Teil meiner Forschung, für welchen ich zunächst die angewandten Methoden sowie die Gründe für deren Wahl darlegen möchte. Grundsätzlich handelt es sich bei der Forschungsmethode um einen qualitativen Ansatz.

Die Wahl eines qualitativen Zugangs anstelle der Anwendung einer quantitativen Methode hat im Wesentlichen zwei Gründe. Erstens ist die Forschung eher induktiv angelegt, und aufgrund ihrer Natur als deskriptive Informationsgewinnung eignet sich der flexible Zugang der qualitativen Forschung besser zum Erkenntnisgewinn. Qualitative Forschungen basieren auf einem sensibilisierenden Konzept, das bedeutet, dass Vorwissen und daraus generierte Theorien natürlich in die Gestaltung der Forschung einfließen. Allerdings geht es nicht darum, bestimmte Hypothesen zu stärken oder zu widerlegen, sondern es sollen laufend aus dem gewonnenen Datenmaterial neue Hypothesen beziehungsweise Informationen gewonnen werden. Dieses Prinzip wird als „Prinzip der Offenheit“ bezeichnet und ist ein wichtiger Vorteil der qualitativen Forschung (Lamnek 2010:19f). Natürlich kommt kein Forschungszugang ohne gewisse Vorannahmen aus. In diesem Fall generieren sich diese aus der bestehenden Forschung zum Respeaking aus anderen Ländern sowie aus den technischen Informationen aus der im vorherigen Kapitel beschriebenen japanischen Forschung.

Darüber hinaus lässt sich ein qualitativer Zugang auch besser an die gegebene Forschungssituation und ihre Limitationen aufgrund von Zeit und örtlicher Distanz sowie Verfügbarkeit von Kontaktpersonen anpassen. Für eine qualitative Forschung ist kein großer Datensatz notwendig, da zum einen keine allgemeingültigen Aussagen wie etwa in den Naturwissenschaften getroffen werden sollen, und zum anderen durch den induktiven Zugang aus dem Einzelfall heraus eine Theorienbildung erfolgen soll. Bei einem bisher noch gar nicht oder wenig beforschten Gebiet, wie es in dieser Arbeit der Fall ist, empfiehlt sich außerdem, die Forschung explorativ anzulegen und ohne vorgefertigte Theorien an das Thema heranzugehen. Da bei meiner Forschung beim Fernsehsender NHK nur eine eingeschränkte Gruppe von befragbaren Personen zur Verfügung stand, war es klar, dass nur eine qualitative Methode zur Anwendung kommen konnte (Lamnek 2010).

Als Methode für die qualitative empirische Forschung wurde das qualitative Interview gewählt. Die Gründe für diese Wahl und der theoretische Hintergrund werden im nächsten Unterkapitel näher erläutert.

5.1 Qualitative Interviews

Das qualitative Interview ist eine beliebte Methode in der qualitativen Forschung, da diese Befragungen relativ einfach und kostengünstig durchzuführen sind. Laut Lamnek sind die Vorteile dieser Methode, dass „die Informationen in statu nascendi aufgezeichnet werden können, unverzerrt-authentisch sind, intersubjektiv nachvollzogen und beliebig reproduziert werden können“ (Lamnek 2010:301). Im Gegensatz zu quantitativen Befragungen zeichnen sich qualitative Interviews dadurch aus, dass sie keine hohe Standardisierung aufweisen, normalerweise mündlich geführt werden, aus offenen Fragen bestehen und der Kommunikationsstil neutral bis weich sein kann (Lamnek 2010).

Interviews lassen sich in verschiedene Unterkategorien einteilen. Dieser Kategorisierung kann aufgrund von unterschiedlichen Aspekten erfolgen. Zum einen spielt die Intention der Befragung eine Rolle. Hier unterscheidet man zwischen vermittelnden und ermittelnden Interviews. Bei einem vermittelnden Interview soll nicht nur Information ausgetauscht werden, sondern der oder die Befragte soll aus dieser Reflektion eine Erkenntnis gewinnen. Diese Art des Interviews eignet sich natürlich nicht dazu, für eine Forschung Informationen zu generieren. Ist dies das Ziel einer Befragung, so spricht man von einem ermittelnden Interview. Im konkreten Fall dieser Forschung handelt es sich am ehesten um ein informatorisches Interview nach Lamnek, da dieses einer deskriptiven Erfassung von Tatsachen dient, wobei die Befragten zumindest teilweise als ExpertInnen verstanden werden.

Ein weiteres Merkmal ist der Grad der Standardisierung. Wie zuvor schon erwähnt wurde, ist ein qualitatives Interview im Vergleich zu einem quantitativen eher wenig standardisiert. Es gibt allerdings innerhalb qualitativer Interviews auch verschiedene Ebenen der Standardisierung, von sehr offen bis hin zu teilstandardisiert. Ein teilstandardisiertes Interview versucht eine Brücke zwischen den beiden Extremen zu bilden. Hierbei sollen die Vorteile beider Arten von Interviews aufgegriffen werden. Durch die niedrigere Standardisierung bleibt ein hohes Maß an Offenheit und Flexibilität erhalten. Zugleich wird durch einen gewissen Grad an vorhandener Standardisierung aber vermieden, dass das Interview vom Thema abschweift. Um mir ebendiese Vorteile der Methode zunutze zu machen, wurde für diese Arbeit ein teilstandardisiertes Interview gewählt. (Lamnek 2010:307ff)

Diese Art der Befragung wird häufig auch als Leitfaden-Interview bezeichnet. Bei diesem Interview wird zwar ein Fragenkatalog zusammengestellt, dieser gilt aber, wie der Name schon sagt, nur als eine Art Leitfaden und dem/der Befragten wird sehr viel eigene Entscheidungsfreiheit in Bezug auf die Beantwortung gelassen. So gibt es hier keine

vorgegebenen Antwort-Kategorien, und der Ablauf des Fragebogens muss bezüglich der Reihenfolge der Fragen nicht eingehalten werden. Bei einem teilstrukturierten Interview gibt es also einen Fragebogen, dieser gibt aber keine Antwortmöglichkeiten vor, sondern besteht nur aus offenen Fragen. Es ergibt sich aber nicht nur für die befragte Person ein relativ großes Ausmaß an Freiheit während der Befragung. Der/die InterviewerIn ist ebenfalls frei, bei Bedarf Nachfragen zu den Aussagen des Gegenübers zu stellen, und kann sogar eventuell im Fragebogen überhaupt nicht erwähnte Themen aufgreifen, wenn diese angesprochen werden. Gerade diese Möglichkeit, durch den Befragten neue Themen einzubringen, ist sehr geeignet für den induktiven Zugang dieser Studie (Hopf 2019:177). In diesem bestimmten Fall ist der Interview-Leitfaden allerdings relativ konkret ausgefallen, da zu großes Abschweifen vom eigentlichen Gegenstand verhindert werden sollte und auf bestimmte Aspekte auf jeden Fall eingegangen werden sollte. Innerhalb des teilstrukturierten Interview-Spektrums kann dieses also als eher strukturiert eingestuft werden.

Teilstrukturierte Interviews können auch noch enger in Typen unterteilt werden, von welchen einer das problemzentrierte Interview ist. Diesem Typ ist meine Umfrage aufgrund der Fokussierung auf ein bestimmtes Thema am ehesten zuzuordnen. Das problemzentrierte Interview stellt laut Hopf einen Kompromiss zwischen einer Orientierung gebenden Leitfaden und der freien Artikulation der InterviewpartnerInnen dar. Das ist auch in Bezug auf die in diesem Fall durchgeführte Befragung relevant (Hopf 2019:177f).

Ein weiteres Merkmal, nach welchem Interviews unterschieden werden können, ist das Medium, über das sie durchgeführt werden. Eine Problematik bei der Durchführung der Interviews ergab sich aber aus der konkreten Situation der für diese Arbeit durchgeführten Befragung. Durch die Restriktionen während der Covid-19-Pandemie war es mir nicht möglich, Befragungen persönlich durchzuführen. Lamnek schreibt in seinem Buch „Qualitative Sozialforschung“, dass sich Telefoninterviews nicht wirklich für qualitative Forschung eignen, da sie den persönlichen Aspekt durch fehlende visuelle Aspekte der Kommunikation verhindern würden (Lamnek 2010:315). Allerdings besteht heutzutage immerhin die Möglichkeit, Interviews über ein Video-Telefonat durchzuführen, was den visuellen Aspekt wieder miteinbringt. Dies wurde auch bei den Interviews mit UntertitlerInnen so durchgeführt. Problematischer waren die Zielgruppeninterviews. Aufgrund der Tatsache, dass die meisten Befragten Schwierigkeiten hätten, einer mündlichen Befragung in japanischer Sprache, insbesondere mittels Videotelefonie, zu folgen und ich der japanischen Gebärdensprache nicht mächtig bin, wurde diese Form der Befragung schließlich für die Forschung ausgeschlossen. Obwohl eine qualitativ angelegte Befragung normalerweise nicht schriftlich erfolgt, war eine

schriftliche Beantwortung der offenen Fragen in diesem Fall die beste Möglichkeit, auch die Meinung der Zielgruppe in diese Arbeit miteinbeziehen zu können. Damit bewegte sich dieser Teil der Forschung dann zwischen der qualitativen und der quantitativen Forschung.

Abgesehen von diesen bisher aufgezählten Merkmalen können sich Interviews auch durch die befragten Personen beziehungsweise das Forschungsinteresse klassifizieren lassen. Weiter oben wurde bereits der Begriff *ExpertIn* benutzt. Auch wenn dieser Begriff ein wenig schwierig zu definieren ist, hat er zumindest für einen Teil der für diese Forschung durchgeführten Interviews Relevanz. ExpertInneninterviews sind in der qualitativen Forschung eine beliebte Form des Interviews. Die Wahl erfolgt häufig aus recht pragmatischen Gründen. ExpertInnen sind oft einfach für Interviews zu motivieren, sie haben einen breiten Wissensschatz und sind eine Art zentrale Stelle, die die Befragung vieler verschiedener Personen unnötig macht. In einem strikt qualitativen Paradigma sind ExpertInneninterviews als qualitative Methode teilweise ein wenig umstritten, da durch die relativ starke Strukturierung und Themeneingrenzung die Prinzipien der Offenheit und der Nicht-Beeinflussung häufig nicht mehr genau eingehalten werden können (Bogner et al. 2014:2f). Da es sich aber bei meiner Methode um eine eher standardisierte Form des qualitativen Interviews handelt, lässt sich die Methode des ExpertInneninterviews gut integrieren.

Eine weitere Debatte, die in der Literatur bezüglich der ExpertInneninterviews geführt wird, betrifft den Begriff des/der ExpertIn als solcher. Schließlich kann, wenn man den Begriff weit fasst, jeder Mensch als ExpertIn auf einem Gebiet aufgefasst werden. Versuche, ExpertInnen nur über das Wissen zu definieren, scheitern hier leicht. Bogner, Littig und Menz führen folgende Definition ein:

Experten lassen sich als Personen verstehen, die sich – ausgehend von einem spezifischen Praxis- oder Erfahrungswissen, das sich auf einen klar begrenzbaren Problemkreis bezieht – die Möglichkeit geschaffen haben, mit ihren Deutungen das konkrete Handlungsfeld sinnhaft und handlungsleitend für Andere zu strukturieren. (Bogner et al. 2014:13)

Damit geht es bei ExpertInnen erstens um deren Einfluss auf die Handlungen anderer und die Struktur, die sie vorgeben, auf der anderen Seite ist aber auch die Intention des Interviews wichtig. Dies liegt daran, dass man die ExpertInnen nicht als Privatpersonen, sondern als in dieser Rolle tätige Menschen versteht. So dient das ExpertInneninterview in erster Linie dazu, Fakten zu einem Sachverhalt zu erfragen. Selbstverständlich spielen auch die Einschätzung und die Erfahrung eines/einer ExpertIn eine Rolle. Es geht jedenfalls nicht darum, Gefühlssituationen oder narrative Erlebnisse widerzuspiegeln (Bogner et al. 2014). Somit fällt der erste Teil meiner Befragung, bei dem UntertitlerInnen befragt werden, durchaus in die

Kategorie ExpertInneninterview, auch wenn es die Zielgruppeninterviews eher nicht tun. Auf diese treffen aber dennoch die anderen, in diesem Kapitel beschriebenen Merkmale zu.

5.2 Inhaltsanalyse

Nachdem das Sammeln der Daten abgeschlossen ist, müssen diese natürlich auch analysiert und dadurch ausgewertet werden, wenn aus ihnen wissenschaftlich verwertbare Erkenntnisse gezogen werden sollen. Obgleich die qualitative Forschung nicht denselben Kriterien der Standardisierung folgen muss wie eine quantitative Methodik und die Auswertung daher im Allgemeinen auch weniger numerisch orientiert ist, muss dennoch eine klar definierte und intersubjektiv nachvollziehbare Methode zur Datenanalyse verwendet werden. Aufgrund der insbesondere in der Sozialforschung steigenden Beliebtheit der qualitativen Methoden gibt es mittlerweile eine Vielzahl an Literatur und Anleitungen zu diesem Thema. Ein gerne genutztes Standardwerk zu Thema ist die qualitative Inhaltsanalyse nach Mayring (2010), auf welcher auch die Analyse der für diese Arbeit gesammelten Daten basiert. Im Folgenden wird diese Form der Inhaltsanalyse näher beschrieben und ihre Anwendung auf die Analyse meiner Interviews detailliert.

Laut Mayring (2010) lässt sich die qualitative Inhaltsanalyse grundlegend durch sechs Punkte definieren: Gegenstand der Analyse ist eine Form der Kommunikation und diese ist in irgendeiner Form festgehalten, zum Beispiel als Protokoll oder Transkription. Die Analyse erfolgt systematisch, wodurch sie sich von anderen Inhaltsanalysen wie der Hermeneutik abgrenzt, und sie läuft nach expliziten Regeln ab, die eine intersubjektive Nachprüfbarkeit möglich machen. Die Analyse geht von einer bestimmten Theorie geleitet vor, und gibt nicht einfach einen Text wieder, sondern bearbeitet ihn. Außerdem analysiert sie das Material nicht als Selbstzweck, sondern um auf bestimmte Aspekte der Kommunikation hinzuweisen und daraus wiederum Rückschlüsse zu ziehen. Abschließend lässt sich auch noch sagen, dass die qualitative Inhaltsanalyse häufig der Hypothesenfindung dient. (Mayring 2010:12f)

Die Inhaltsanalyse nach Mayring basiert auf zuvor entstandenen Methoden der Analyse wie der *content analysis*, der Hermeneutik, des interpretativen Paradigmas und der Psychologie der Textverarbeitung. Dabei werden die Vorteile dieser Methoden übernommen und weniger hilfreiche Aspekte weggelassen. So wurden zum Beispiel aus der *content analysis* die systematische Vorgehensweise und Gütekriterien übernommen. Aus der Hermeneutik wurde übernommen, dass eine intensive Quellenkritik bezüglich der Authentizität des jeweiligen Textes durchgeführt werden muss und außerdem die Vorannahmen des Forschenden klar

dargelegt werden sollen. Aus diesen Methoden sowie der Miteinbeziehung quantitativer Zugänge bilden sich folgende Kriterien heraus, denen eine Inhaltsanalyse laut Mayring folgen muss:

- Das Material muss in den Kommunikationszusammenhang eingebettet werden, das heißt, das Material kann nicht ohne seinen Kontext gesehen werden.
- Das Vorgehen muss systematisch und regelgeleitet sein. Zur Analyse ist die Bildung von Kodierungseinheiten und Kategorien notwendig.
- Das Kategoriensystem ist das zentrale Instrument der Analyse.
- Obwohl die Analyse strukturiert erfolgt, muss sie dennoch an den Gegenstand angepasst werden, und nicht blind den Regeln folgen.
- Verfahren sollen in einer Pilotstudie getestet werden.
- Quantitative Analyseschritte sollen miteinbezogen werden, zum Beispiel indem man wiedergibt, wie oft eine Kategorie auftritt.
- Die Gütekriterien Objektivität, Reliabilität und Validität müssen erfüllt sein. (Mayring 2010: 48-51)

5.3 Ablauf der Inhaltsanalyse

In diesem Unterkapitel wird beschrieben, wie eine Inhaltsanalyse laut Mayring generell ablaufen soll, und es wird zugleich ein Bezug zu der in dieser Arbeit durchgeführten Analyse hergestellt.

Als erster Schritt muss selbstverständlich festgelegt werden, welches Material überhaupt Teil der Analyse werden soll. Im Falle dieser Arbeit mussten keine Stichproben aus den gewonnenen Aufzeichnungen genommen werden, sondern es konnte das gesamte Material, also die Transkripte aller drei mündlich durchgeführten und die Texte aus den beiden schriftlichen Befragungen, miteinfließen. Anschließend müssen der Hintergrund und die Entstehung des Materials sowie die daran beteiligten Personen dargelegt werden (Mayring 2010:52). Dies geschieht im nächsten Unterkapitel, das sich mit der Suche nach InterviewpartnerInnen befasst.

Als nächster Schritt befasst sich der oder die Forschende mit der Fragestellung, um die Richtung der Analyse festzulegen. Hierbei kann Material auf verschiedene Arten analysiert werden, auf einer inhaltlichen sowie metatextuellen Ebene. Die Analyse muss also von der dahinterstehenden Theorie geleitet werden. Im Falle dieser Forschung war die Fragestellung relativ weit gefasst, weswegen verschiedene Bereiche aus der Untertitelung miteinfließen

können. Allerdings gibt es dennoch Gesprächsteile, die für die Beantwortung der Frage nicht relevant sind, so wurde zum Beispiel an ein paar Stellen die Untertitelung in Österreich besprochen. Die Analyse soll im Großen und Ganzen die Frage beantworten, wie Untertitelung in Japan stattfindet, welche Probleme dabei auftreten und wie die Meinung der Hauptzielgruppe zu den so produzierten Untertiteln ist (Mayring 2010:56f).

Im Anschluss an diese Schritte beginnt die eigentliche Analyse. Hierzu müssen als erstes Analyseeinheiten festgelegt werden. Hierbei wird die Kodiereinheit festgelegt, dabei handelt es sich um den kleinsten zu untersuchenden Textabschnitt, die Kontexteinheit, also den größten miteinzubeziehenden Abschnitt, und die Auswertungseinheit, die festlegt, welche Textteile nacheinander ausgewertet werden. Im Fall dieser Arbeit handelt es sich bei der kleinsten Einheit um eine sinnhafte Aussage, also eine Proposition. Die Kontexteinheit ist das gesamte durchgeführte Interview. Eine Auswertungseinheit wird thematisch abhängig gemacht; die Auswertung erfolgt im Fall dieser Forschung hauptsächlich chronologisch, da auch in den Interviews die zu behandelnden Themengebiete meistens chronologisch abgearbeitet werden. Darüber hinaus werden die anzuwendenden Kategorien und Kodizes festgelegt, die im zu analysierenden Text dann markiert werden. Auch hierzu dienen wieder die festgelegten Themengebiete beziehungsweise in manchen Fällen auch Unterfragen zu den Themengebieten, wie zum Beispiel die Frage nach der Überprüfung von Qualität (Mayring 2010:59). Alle in der Analyse verwendeten Kodierungseinheiten in dieser Beschreibung aufzulisten, wäre aus Platzgründen nicht sinnvoll, jedoch werden sich diese in der anschließenden Analyseerläuterung der Befragung widerspiegeln.

Nach einem ersten Herausfiltern der im Text relevanten Stellen und dem Einordnen in Kategorien folgt die eigentliche Analyse des Textes. Diese soll nach drei grundlegenden qualitativen Methoden erfolgen, der Zusammenfassung, der Explikation und der Strukturierung. Bei der Zusammenfassung geht es darum, das Material auf eine überschaubare Menge zu reduzieren, wobei trotzdem noch die wichtigen Kernaussagen wiedergeben werden sollten. Im Laufe der Zusammenfassung erfolgt außerdem eine weitere, induktive Kategorienbildung, da die behandelten Kategorien bei der Verkürzung augenfällig werden. Bei der Explikation wird zusätzliches Material herangetragen, das das Verständnis der Textstelle erweitert und diese außerdem deutet und erklärt. Im Fall dieser Arbeit fließen hier natürlich auch die Erkenntnisse aus den vorherigen Theoriekapiteln erneut ein. Schließlich folgt noch ein Schritt der Strukturierung, bei dem das Material anhand von zuvor festgelegten Kriterien eingeordnet wird und ein Querschnitt hindurch gezogen wird. Dadurch werden die Zusammenhänge noch einmal verständlicher gemacht. Diese Strukturierung kann auf einer formalen, einer inhaltlichen, einer

typisierenden oder skalierenden Ebenen erfolgen. Natürlich gibt es auch noch wesentlich tiefer gehende Analyseformen, diese sind für die vorliegende Arbeit aber nicht notwendig (Mayring 2010:66).

Die auf diese Art durchgeführte Analyse wird dann im sechsten Kapitel dieser Arbeit verschriftlicht. Der nächste Abschnitt befasst sich mit der Durchführung der Forschung, beginnend mit der Suche nach InterviewpartnerInnen.

5.4 Durchführung der Forschung

In diesem Kapitel wird der tatsächliche Ablauf der für diese Arbeit durchgeführten Forschung inklusive der dabei aufgetretenen Schwierigkeiten beschrieben, um die Hintergründe der gewonnenen Informationen und die darin involvierten Personen transparent zu machen.

5.5 Die Suche nach InterviewpartnerInnen

Wie bereits besprochen, behandelt die auf Japanisch veröffentlichte Literatur zum Thema Respeaking beim Fernsehen praktisch ausschließlich die Anwendung bei NHK. Im Zuge meiner Recherche stellte sich heraus, dass durchaus auch andere Fernsehsender bei ihrer Untertitelung mit Respeaking arbeiten und Fremdfirmen mit dieser beauftragen. Dass NHK so stark vertreten ist, liegt wohl auch daran, dass der Anbieter ein eigenes Forschungszentrum betreibt. Jedenfalls ist Respeaking als Methode auch auf unterschiedlichen Webseiten des Senders NHK repräsentiert, daher sah ich NHK als meine erste Anlaufstelle an.

Die erste Kontaktaufnahme erfolgte über ein Anfrageformular auf der Seite der NHK-Subfirma NHK Global Media Services, die das Untertitelungszentrum von NHK betreibt. In dieser Anfrage erkundigte ich mich nach der Möglichkeit, Informationen zu der Live-Untertitelung mittels der Respeaking-Methode zu erhalten und eventuell zu diesem Thema auch Interviews zu führen. Durch den zuständigen Produzenten wurde ein Interview mit den RespeakerInnen zugesagt. Bis zur tatsächlichen Durchführung der Interviews änderte sich die Situation jedoch wieder und die zuständigen Personen zogen ihre Erlaubnis, mit RespeakerInnen direkt zu sprechen, leider zurück. Ich führte allerdings Ende März ein Interview mit oben genanntem Produzenten und einem Regisseur. Letzterer war, wie sich herausstellte, eigentlich als Dolmetscher engagiert worden und sprach bei dem Interview kaum. Behandelt wurden in diesem Interview allgemeine Fragen zur Untertitelung mit Respeaking sowie Themen wie Feedback, Ausbildung und Fähigkeiten der RespeakerInnen. Darüber hinaus

bot mir der Produzent Hilfe dabei an, doch noch Interviews mit RespeakerInnen führen zu können.

Wie sich herausstellte, beauftragt NHK für Live-Untertitelung besonders im Bereich der Nachrichtensendungen eine Fremdfirma mit dem Namen Faith, die ebenfalls mit Respeaking arbeitet. Durch den Produzenten von NHK bekam ich den Kontakt zum CEO von Faith. Dieser erklärte sich sofort zu einem Interview bereit, das noch am nächsten Tag durchgeführt wurde. Gesprochen wurde über die Geschichte seiner Firma, die seit 2014 Live-Untertitelung mit Respeaking anbietet, das Verhältnis zu NHK, die Untertitelungsmethode, die etwas anders ist als die bei NHK selbst eingesetzte, Qualitätsstandards und die Ausbildung. Auch zeigte er mir via Webcam das Aufnahmestudio, in dem die RespeakerInnen arbeiten. Es gelang mir dann auch, mit einem der Respeaker direkt zu sprechen, dies fand in einem separaten Videogespräch statt. Alle Interviews wurden mittels Videotelefonie durchgeführt und über Audioaufnahme aufgezeichnet.

Für eine umfassende Darstellung der Untertitelung in Japan ist es allerdings wichtig, auch die Zielgruppe miteinzubeziehen. Wie im Theorieteil bereits ausführlich diskutiert wurde, ist die Gruppe, die Untertitelung verwendet, eigentlich sehr breit. Dennoch werden die intralingualen Untertitel primär mit der Zielsetzung, Gehörlosen und Schwerhörigen einen Zugang zum Medium Fernsehen zu erlauben, erstellt. Daher entschied ich mich, zu versuchen, Meinungen dieser Zielgruppe über die Qualität und die Probleme der im Fernsehen ausgestrahlten Untertitel zu erfahren. Hierbei tat sich aber ein logistisches Problem auf. Meine bisherigen Interviews wurden über eine Webcam in Lautsprache durchgeführt, eine Form der Kommunikation, die für die meisten Mitglieder dieser Zielgruppe eine Unmöglichkeit darstellt. Daher war es notwendig, einen anderen Modus für diese Interviews zu finden.

Nach einigen Abwägungen entschied ich mich, den Zielpersonen einen Fragebogen mit offen formulierten Fragen schriftlich zukommen zu lassen. Dies stellt eine Abweichung von der gewöhnlichen Methodik eines qualitativen Interviews dar, deren Konsequenzen mitbedacht werden müssen. Der Zweck des zuvor angewandten teilstrukturierten Interviews ist schließlich auch, dass der/die InterviewerIn die Möglichkeit bekommt, durch Nachfragen und Ergänzungen vom „Programm“ abzuweichen und sich an die Situation anzupassen. Durch eine schriftliche Übermittlung des Fragebogens gibt die interviewende Person also auch einen Teil der Kontrolle über den Ablauf des eigenen Interviews aus der Hand. Das Interview ist dadurch standardisierter, da alle Befragten in jedem Fall dieselben Fragen beantworten. Da diese Fragen aber immer noch offen gestellt sind, steigt natürlich auch die Wahrscheinlichkeit, dass Missverständnisse auftreten können. Schließlich können die befragten Personen Fragen

unterschiedlich auffassen. Eine Möglichkeit, Variablen in der Beantwortung zu reduzieren, wäre, statt offener Fragen geschlossene Fragen zu nutzen und somit mehr auf die Techniken des quantitativen Interviews zuzugreifen. Da aber keine ausreichende Anzahl von zu befragenden Personen erreicht werden konnte und diese Methode für die Art der Fragen, die es zu stellen galt, auch nicht als ideal geeignet erschien, fiel die Wahl schließlich doch auf einen offenen Fragebogen. Wären die Ressourcen dazu verfügbar gewesen, wäre die ideale Methode wohl eine Befragung unter Zuhilfenahme eines/r DolmetscherIn gewesen, was aber im Rahmen der Masterarbeit nicht möglich war.

Nach der Klärung der Frage nach der geeigneten Methodik begann die Suche nach InterviewpartnerInnen. Der Kontakt kam hierbei über zwei verschiedene Gehörlosenorganisationen zustande.

Die erste Organisation ist die 全日本ろうあ連盟, auf Englisch als *Japanese Federation of the Deaf* (kurz JFD) übersetzt. Die JFD bildet den Dachverband aller in den Japanischen Bezirken lokal aktiven Organisationen für Gehörlose. Sie definiert ihre Zielsetzung darin, die Rechte von Gehörlosen zu stärken und ihre Wohlfahrt zu fördern. Hierbei hebt sie insbesondere folgende Punkte hervor: die Institutionalisierung und Anerkennung von Gebärdensprach-DolmetscherInnen, die Beseitigung von Diskriminierung auf Basis einer Hörbehinderung und die Förderung der selbständigen Teilnahme hörbehinderter Personen in der Gesellschaft (Japanese Federation of the Deaf 2021). Die Kontaktaufnahme fand über ein Kontaktformular auf der Webseite der Organisation statt, in dem ich die Zielsetzung meiner Forschung und die Suche nach Kontaktpersonen beschrieb. Im anschließenden Emailverkehr wurde ich von einem Vertreter darauf aufmerksam gemacht, dass die meisten dort tätigen Personen Gebärdensprache als tägliche Kommunikationsmethode nutzen und normalerweise einen hohen Grad an Schwerhörigkeit aufweisen. Um möglicherweise auch Personen mit einer geringeren Höreinschränkung oder anderer Kommunikationssprache als der Japanischen Gebärdensprache zu erreichen, wurde mir der Kontakt zu einer anderen Organisation vermittelt.

Bei dieser Organisation handelte es sich um die 全日本難聴者・中途失聴者団体連合会 [Zennihon nanchō-sha chūtooshitchō-sha dantai rengō-kai (kurz: Zennancho)]. Der Name der Organisation heißt übersetzt „Organisation der Personen mit Hörschädigungen und Personen mit post-linguaalem Hörverlust in ganz Japan“ und richtet sich somit vor allem auch an jene, die erst nach dem Spracherwerb ihr Gehör zu verlieren begonnen haben. Die Zennancho hat es sich zum Ziel gesetzt, den gesellschaftlichen Status von hörbehinderten Personen zu verbessern sowie ihre Wohlfahrt und Teilnahme an der Gesellschaft zu fördern. Zu diesem Zweck führt sie auch wissenschaftliche Untersuchungen durch. Wie der Name

bereits vermuten lässt, ist sie ebenfalls in ganz Japan aktiv (Zennancho 2021). Beide Verbände zeigten sich sofort bereit, bei meiner Forschung behilflich zu sein. Es wurden schwerhörige und gehörlose Personen aus beiden Organisationen befragt, wobei sich durch die Verbreitung über eine Vermittlungsperson leichte Unterschiede im Ablauf der Befragung ergaben, die in der Auswertung näher erklärt werden. Zunächst soll jedoch der Aufbau der beiden Fragebögen im nächsten Kapitel beschrieben werden.

5.6 Fragebogaufbau

Bei der Befragung der Zielpersonen kommen zwei verschiedene Fragebögen zum Einsatz. Die erste Variante befasst sich mit Fragen der Erstellung von Live-Untertiteln insbesondere mit der Respeaking-Methode und richtet sich somit vor allem an die Untertitel-RedakteurInnen und RespeakerInnen. Der zweite Fragebogen ist hingegen an die Zielgruppe gerichtet. Der erste Fragebogen orientiert sich an jenen Themen, die in den im Theorieteil besprochenen Arbeiten erwähnt wurden. Der zweite Fragebogen hat seine Basis in den Befragungen von gehörlosen und schwerhörigen NutzerInnen in verschiedenen europäischen Ländern, wenngleich diese Arbeit natürlich nicht beansprucht, dieselbe Reichweite oder einen ähnlichen Umfang zu besitzen. Beide Fragebögen wurden auf Deutsch verfasst und anschließend ins Japanische übersetzt, um sie in dieser Arbeit in beiden Versionen zur Verfügung stellen zu können. Die japanischen Versionen der Fragebögen wurden von einem Muttersprachler Korrektur gelesen, um eventuelle Missverständnisse aufgrund der Fremdsprache zu vermeiden. Im folgenden Unterkapitel wird der Aufbau der Fragebögen beschrieben und die einzelnen Fragen werden näher erklärt.

5.6.1 Fragebogen für NHK

Der Fragebogen, der sich an den Fernsehsender NHK richtet, besteht aus zwei klar unterschiedlichen Teilen. Einer behandelt allgemeine Fragen zur bei NHK eingesetzten Untertitelung, und kann auch von jemandem beantwortet werden, der oder die keine persönliche Erfahrung mit Untertitelung beziehungsweise dem Einsatz der Respeaking-Methode hat. Hier geht es darum, Informationen dazu zu bekommen, wie de facto momentan bei NHK Untertitelt wird. Im zweiten Teil, der sich an Personen richtet, die tatsächliche Erfahrung mit Respeaking haben, soll hingegen ihre persönlichere Sichtweise auf ihre Arbeit in Erfahrung gebracht werden.

Am Anfang des ersten Teils des Fragebogens stehen Fragen zur Geschichte, Entwicklung und momentanen Situation der Live-Untertitel bei NHK. Diese sollen klären, seit wann Live-Untertitelung bei NHK eingesetzt wird, welche Methoden verwendet werden und wie viel Prozent der ausgestrahlten Sendungen zum jetzigen Zeitpunkt untertitelt werden. Da Respeaking ja der Hauptfokus dieser Arbeit ist, wird außerdem die Frage gestellt, für welche Sendungen speziell diese Methode eingesetzt wird. Des Weiteren ist interessant, ob nur intralingual im Japanischen oder auch interlingual in andere Sprachen untertitelt wird, beziehungsweise ob es Pläne gibt, zu einem späteren Zeitpunkt die interlinguale Untertitelung auch einzusetzen. Anschließend beschäftigt sich der Fragebogen mit der Rollenaufteilung und Arbeitsweise der RespeakerInnen und EditorInnen, die, wie aus dem Theorieteil bekannt, beim japanischen Fernsehen zwei getrennte Rollen innehaben. Ebenso beschäftigt sich eine Frage mit den Kriterien, die berücksichtigt werden müssen, wenn neue RespeakerInnen eingestellt werden.

Der zweite Teil der allgemeinen Fragebogenhälfte befasst sich mit Qualitätskriterien. Zunächst soll untersucht werden, wie Qualität bei der Untertitelung überprüft wird und ob es bestimmte Richtlinien gibt, die dabei zum Einsatz kommen. Auch Formatvorschriften und Zielsetzungen bezüglich der maximalen Verzögerung der Untertitel sollen erfragt werden. Abgeschlossen wird diese Sektion dann mit einer Frage zu editierten und wörtlichen Untertiteln. Hierbei soll sowohl herausgefunden werden, ob mehr Wert auf Wörtlichkeit oder Korrektheit gelegt wird, als auch ob eine schnellere Ausstrahlung der Untertitel wichtiger als die Wörtlichkeit derselben ist.

Der dritte Teil schließlich beschäftigt sich mit der Zukunft der Untertitelung bei NHK und den Möglichkeiten, Feedback von Zielgruppen zu erhalten. Hierbei soll erfragt werden, inwieweit NHK mit seinen Hauptzielgruppen, insbesondere jener der gehörlosen und schwerhörigen Personen, in Kontakt steht. Außerdem wird in Erfahrung gebracht, welche Weiterentwicklungen für die Untertitelung geplant sind und welche Zukunft sich die befragten Personen dafür vorstellen können.

Im Anschluss an den allgemeinen Teil folgen die Fragen an die RespeakerInnen. Dieser beginnt mit Fragen zu dem Hintergrund der RespeakerInnen, ihrer Ausbildung und wie sie die Arbeit eines Respeakers gelernt haben. Anschließend werden Fragen zur Arbeitsweise gestellt, wie häufig und für welche Sendungen sie mit dem Respeaking arbeiten und ob sie neben dem Respeaking beziehungsweise der Untertitelung andere Aufgaben haben. Darüber hinaus wird eine Frage nach ihren allgemeinen Arbeitszeiten gestellt, um unter anderem herauszufinden, ob es sich hier um Vollzeit- oder Teilzeit-Angestellte handelt.

Nach diesen Informationen zur Arbeitsweise beschäftigt sich der Fragebogen mit den auftretenden Problematiken und Lösungsansätzen beim Respeaking. Auch nach dem Umgang mit Stress und besonders schwierigen Situationen wird gefragt. Die nächste Frage beschäftigt sich mit einer Problematik, die auf eine Eigenheit der japanischen Sprache zurückzuführen ist. Das Japanische weist grundsätzlich relativ viele Homonyme auf. Diese Homonyme werden durch unterschiedliche Schriftzeichen-Kombinationen wiedergegeben und auch unterschieden. Das Spracherkennungsprogramm kann aber natürlich nicht immer aus dem Kontext heraus die gemeinte Wortbedeutung und damit die korrekte Zeichenkombination unterscheiden. Daher ist es von Interesse, bis zu welchem Grad das Programm diese Unterscheidung überhaupt treffen kann und an welchem Punkt menschliche EditorInnen eingreifen müssen.

Gegenstand einer Frage ist ebenfalls die Wörtlichkeit der Untertitel bezüglich derer die TeilnehmerInnen um eine Einschätzung gebeten werden. Auch wird mit den Befragten diskutiert, welche Eigenschaften AnwarterInnen mitbringen müssen, um gute RespeakerInnen zu sein. Diese Qualifikationen sollen auch mit jenen von DolmetscherInnen verglichen werden, um zu sehen, ob die Befragten eine Beziehung zwischen diesen beiden Professionen herstellen können.

5.6.2 Fragenbogen für Zielgruppe

Obgleich die Befragten die Möglichkeit bekamen anonym an der Befragung teilzunehmen, werden am Anfang des Fragebogens einige persönliche Informationen abgefragt, wie das Geschlecht, das Alter und die berufliche Tätigkeit der Person. Diese Informationen sollten dazu dienen, eine Vorstellung davon zu bekommen, wer die befragten Personen sind. Darüber hinaus werden auch für diese Forschung wesentlich wichtigere Fragen zur Person gestellt, nämlich das Ausmaß des Hörverlustes, das Alter in welchem dieser eingetreten ist, ob sie Hörgeräte nutzen und welche Sprache sie als ihre Muttersprache bezeichnen würden. Diese Informationen haben mit großer Wahrscheinlichkeit eine Auswirkung darauf, wie und in welcher Frequenz sie Programme mit Untertitelung konsumieren, und sind daher für die Forschung relevant.

Der nächste Teil des Fragebogens beschäftigt sich mit dem Fernsehverhalten der interviewten Personen und befragte sie dazu, wie häufig sie untertitelte Fernsehsendungen konsumieren, welche Sendungen sie sehen und ob sie auch bei Live-Sendungen Untertitelung nutzen. Diese Fragen dienen vor allem dazu, einen Überblick über die Fernsehgewohnheiten zu bekommen und bieten außerdem einen einfachen Einstieg in das Interview. Danach folgen Fragen zur Qualität der Untertitel und eventuellen Verbesserungsvorschlägen. Auch wird

gefragt, welche Kriterien generell wichtig für gute Untertitel sind und wie Fernsehen barrierefreier gemacht werden kann.

Der dritte Teil des Fragebogens beschäftigt sich mit detaillierten Fragen zur Ausführung der Untertitel und stellt Fragen zur Geschwindigkeit, zu editierten und wörtlichen Untertiteln, wie ein Sprecherwechsel in den Untertiteln angezeigt werden soll und wie lesbar die Untertitel eingeschätzt werden. Außerdem werden die TeilnehmerInnen gebeten, einen Vergleich anzustellen, ob es Fernsehsender gibt, welche ihrer Meinung nach bessere Untertitel erstellen als andere. Der letzte Teil des Fragebogens versucht noch festzustellen, ob die Befragten wissen, wie Live-Untertitel hergestellt werden und welche Probleme dabei auftreten. Eine offene Frage danach, ob sie noch etwas hinzuzufügen wissen, soll eventuelle andere Punkte, die bei der Fragebogenerstellung nicht bedacht wurden, den TeilnehmerInnen aber als wichtig erscheinen, miteinbeziehen.

6 Ergebnisse

6.1 Qualitative Interviews mit Untertitlern

In den folgenden Kapiteln werden die Ergebnisse der Interviews mit den Vertretern von NHK und Faith erläutert. Die Resultate sind hierbei in die Kategorien „Allgemeine Informationen“, „Fähigkeiten und Arbeitsweise der RespeakerInnen“, „Qualitäts- und formale Kriterien“, „Herausforderungen und Lösungsansätze“, sowie „Zukunft des Respeakings in Japan“ unterteilt. Da sich die Abläufe in den Arbeitsweisen zwischen NHK selbst und der Fremdfirma teilweise deutlich unterscheiden, werden sie einander gegenübergestellt. Weiter sollte noch festgehalten werden, dass es sich hierbei natürlich nur um Informationen bezüglich der in den Abteilungen der interviewten Personen durchgeführten Untertitelung handelt und diese daher nicht den Anspruch erheben, für die Untertitelung in ganz Japan Geltung zu besitzen. Wie man an den Ergebnissen sehen kann, sind die Unterschiede bereits zwischen zwei Arbeitsplätzen relativ groß, und es ist anzunehmen, dass es auch beispielsweise bei anderen Fernsehsendern verschiedene Herangehensweisen gibt.

6.1.1 Allgemeine Informationen zur Untertitelung bei NHK und Faith

In diesem Unterkapitel werden die allgemeinen Arbeitsabläufe, die untertitelten Live-Sendungen und die beteiligten Personen vorgestellt. Einige Informationen wurden bereits in der Zusammenfassung der vorherigen Arbeiten zu dieser Thematik bei NHK erwähnt, jedoch wird mit der Wiedergabe der Interviews die Ist-Situation zutreffender dargestellt und durch die Möglichkeit, nachzufragen, konnten einige der zuvor bestandenen Unklarheiten gelöst werden.

Grundsätzlich werden bei **NHK** Live-Untertitel über zwei verschiedene Kanäle erstellt. Auf der einen Seite gibt es Live-Untertitelung, die *in-house* in den Studios von NHK selbst durchgeführt wird. Für andere Sendungen werden aber auch Fremdfirmen beauftragt. Bei NHK selbst wird die Untertitelung noch einmal nach Art der Sendung eingeteilt. Diese Einteilung unterscheidet zwischen den Nachrichtenuntertiteln (ニュース字幕) und den Sendungsuntertiteln (番組字幕), unter die alle Sendungen fallen, die keine Nachrichten- oder Informationssendungen sind, wie zum Beispiel Sportsendungen oder Talkshows.

Bei der zur Untertitelung verwendeten Software handelt es sich nach wie vor um die bereits in dieser Arbeit beschriebene vom Technischen Forschungszentrum von NHK (NHK 技術研究所) direkt entwickelte Software (Homma et al. 2007, Homma 2010, Satou 2018), obwohl diese über die Jahre einige Updates erfahren hat und in Kürze größere Erneuerungen

geplant sind. Bei NHK sind momentan etwa 98 Prozent des Programmes untertitelt. Dazu, wie viel Prozent der Live-Sendungen genau bereits untertitelt werden, lagen leider keine konkreten Zahlen vor. Allerdings ergibt sich aus der allgemein hohen Untertitelungsrate und den Aussagen der Interviewpartner, dass es sich hier ebenfalls um einen recht hohen Anteil handelt. Dies folgt auch den Richtlinien der Regierung bezüglich der Untertitelung, welche in dieser Arbeit auch bereits besprochen wurden. Live-Sendungen, die nicht durchgehend untertitelt sind, sind insbesondere jene, deren Untertitelung als besonders schwierig erachtet wird – als Beispiele genannt wurden Expertendiskussionen und Verhandlungen im Parlament. Live-untertitelt wird seit den frühen 2000er Jahren, allerdings haben auch die Vertreter von NHK die Erdbebenkatastrophe im März 2011 als großen Impetus herausgestrichen, die Untertitelungsrate weiter zu erhöhen.

Es wird bei NHK allerdings nicht nur Respeaking angewandt, sondern auch andere Methoden. Beispielweise werden Fremdfirmen für die Untertitelung mittels Tastatur beauftragt, und für einige Sendungen wird die Methode der direkten Erkennung genutzt. Immer mit Respeaking untertitelt werden Sendungen, bei denen häufig mehrere Personen zugleich sprechen, wie beispielsweise Talkshows. Auch für Sportsendungen wird Respeaking als Methode genutzt. Generell kommt es jeden Tag zu Situationen, in denen Respeaking genutzt wird.

Es gelang mir, wie zuvor beschrieben, auch ein Interview mit dem CEO von **Faith**, einer der Fremdfirmen, die NHK momentan beauftragt, zu führen. Diese Firma nutzt seit dem Jahr 2014 die Respeaking-Methode, um Live-Untertitelung durchzuführen, am Anfang noch für Sportkanäle, später dann auch für NHK. Sie arbeitet nur mit Spracherkennung, die von RespeakerInnen durchgeführt und durch EditorInnen unterstützt wird. Bereits im Jahr 2012 wurde mit der Forschung zum Thema Live-Untertitelung begonnen. Die Spracherkennungssoftware als Grundprodukt wurde zugekauft und anschließend für die Untertitelung weiterentwickelt. Das so entstandene Produkt heißt „amivoice“ und unterscheidet sich von der bei NHK verwendeten Software. Der Grund, wieso eine andere Software genutzt wird, ist hauptsächlich der Kostenfaktor. Da es sich um ein kleineres und neueres Unternehmen handelt, wäre die Software von NHK einfach zu kostenintensiv. Die Erkennungsrate von amivoice ist allerdings niedriger als jene des von NHK verwendeten Eigenprogramms. Mittlerweile nutzen auch andere Unternehmen das amivoice-Produkt. So verwendet beispielweise der Sportsender J-sports dieselbe Software, führt die Untertitelung aber *in-house* durch, anstatt sie an Faith auszulagern.

Neben NHK produziert das Unternehmen auch noch für andere große Sender, wie zum Beispiel Fuji-TV, Untertitel. Zumeist funktioniert dies so, dass die Untertitelung in den Studios von Faith durchgeführt und dann elektronisch an die Sender übertragen wird. Es gibt außer Faith noch etwa fünf bis sechs andere Firmen, die im Live-Untertitelungsbereich arbeiten, allerdings arbeiten diese nicht alle mit Respeaking, sondern teilweise auch mit Steno-Tastaturen. Etwa 80 Prozent der Untertitelung wird insgesamt mit der Tastatur-Methode durchgeführt, Respeaking ist also in der Live-Untertitelung noch eher nicht so verbreitet.

Es werden auch bei Faith verschiedene Arten von Sendungen untertitelt, wie zum Beispiel Nachrichten und Sportprogramme, aber auch Variety-Sendungen. Für NHK speziell werden aber vor allem Nachrichten- und Informationssendungen untertitelt, unter anderen auch Sendungen, in denen ExpertInnen zu verschiedenen Themen sprechen. Bei diesen „Erklärsendungen“ kommt eine Art der Hybridmethode zum Einsatz, da es ein fixes Skript gibt, aus dem die Untertitel bezogen und eingespielt werden können. Sollte die sprechende Person von diesem abweichen, steht ein(e) RespeakerIn bereit, um fallweise zu übernehmen, anders als z.B. bei Sportsendungen, wo in jedem Fall Respeaking stattfindet.

Anders als bei NHK haben die RespeakerInnen bei Faith keine besondere Spezialisierung auf eine Art von Sendung, sondern untertiteln alle Genres, für die Faith zuständig ist. Es spielt allerdings eine Rolle, wie lange die betreffende Person bereits für Faith arbeitet. Neue MitarbeiterInnen untertiteln eher keine Sendungen, bei denen die Abläufe sehr schnell sind. Mit der Zeit sollen aber alle so eingeschult sein, dass sie im Bedarfsfall jede Sendung übernehmen könnten. Dies ist auch dann wichtig, wenn ein/e MitarbeiterIn beispielsweise aus Krankheitsgründen ausfällt.

Wie aus den zuvor zitierten Arbeiten bereits erkennbar, wird die Untertitelungsarbeit bei **NHK** nach wie vor im Team durchgeführt, wobei normalerweise der oder die RespeakerIn von zwei EditorInnen unterstützt wird. Ein Arbeiten alleine ist laut den Interviewpartnern unmöglich, da durch die hohe Anzahl an Homonymen der japanischen Sprache und andere Eigenheiten zu viele Fehler auftreten würden. Wenn die Korrekturarbeit auf zwei oder mehr Personen aufgeteilt ist, ist jeder für einen Satzteil zuständig, den ihm das Programm automatisch zuteilt. Hierbei hören sie aber natürlich den gesamten Dialog mit, um die Zusammenhänge zu kennen. Diese Art der Aufteilung wird bei NHK seit dem Jahr 2014 so praktiziert. Sowohl die RespeakerInnen als auch die EditorInnen wechseln sich während einer Sendung ab. Daher müssen während einer längeren Sendung auch mehr als ein/e RespeakerIn anwesend sein. Der Wechsel findet etwa alle 15 Minuten statt. Bei den EditorInnen wird in dieser Situation zumeist mit drei Personen gearbeitet, das heißt es findet zwar im selben

Intervall ein Wechsel statt, aber da immer nur eine Person herausgewechselt wird, sind die Zeiten, die eine einzelne Person durchgehend im Einsatz ist, insgesamt ein wenig länger. Zu den EditorInnen und RespeakerInnen ist noch zu bemerken, dass es sich bei NHK hierbei um getrennte Positionen handelt, die RespeakerInnen also nie die Rolle der EditorInnen übernehmen und auch nicht umgekehrt.

Ein Wechsel der RespeakerInnen ist bei **Faith** im Vergleich zu NHK grundsätzlich nicht angedacht. Für den Bedarfsfall ist allerdings immer eine Person anwesend, die theoretisch übernehmen kann. Besonders bei Sendungen für NHK, wo es keine Werbeunterbrechungen gibt, führt dies zu langen Sprechzeiten ohne Pause für die RespeakerInnen. Die KorrektorInnen können pausieren, indem sie dies mit einem Knopf auf ihrem Computer angeben. Ihre Arbeit wird dann kurzzeitig den anwesenden EditorInnen zugeteilt. Es gibt aber auch hier keine genauen Vorgaben dazu, wann Pausen gemacht werden sollten, und somit keinen offiziellen Wechsel.

Bei der Firma Faith können aufgrund der schlechteren Spracherkennungsraten der Software die meisten Sendungen nicht nur mit zwei EditorInnen untertitelt werden, sondern es kommen üblicherweise drei bis vier Personen für die Korrektur zum Einsatz. Wie viele es tatsächlich sind, hängt aber auch von der Art der Sendung ab. Drei oder vier Personen sind für Nachrichtensendungen oder Variety-Sendungen notwendig, bei Sportsendungen reichen allerdings zwei MitarbeiterInnen. Bei Nachrichtensendungen sind deshalb mehr MitarbeiterInnen notwendig, weil die Informationsdichte hoch ist und es außerdem wichtig ist, dass wirklich das gesamte Gesagte untertitelt wird. Variety-Sendungen haben zwar keinen so hohen Anspruch an die Vollständigkeit, dafür ist die Untertitelung aber generell schwieriger, da es keinen fixierten Ablauf gibt und häufig mehrere Personen gleichzeitig sprechen, was die Erkennungsrate des Programmes vermindert. Anders ist dies bei Sportsendungen oder Tele-Shopping, wo das Gesprochene keine so hohe Dichte aufweist oder auch zusammengefasst werden kann. Bei Sportsendungen muss zudem ohnehin öfters zusammengefasst werden, da durch die Verzögerung ansonsten Unterschiede zwischen Bild und Ton entstehen könnten, die verwirrend wären.

Untertitelung wird bei **NHK** momentan nur intralingual durchgeführt und es gibt auch keine konkreten Pläne, in der näheren Zukunft interlinguale Untertitelung anzuwenden. Allerdings gab es bereits ungefähre Überlegungen dazu, diese für die 2020 geplanten Olympischen Spiele in Erwägung zu ziehen. Da diese aufgrund von Covid19 aber verschoben wurden, waren zum Zeitpunkt des Interviews keine weiteren Pläne diesbezüglich bekannt. Auch bei **Faith** wird keine intralinguale Untertitelung durchgeführt, es gäbe aber schon

Überlegungen dazu. Allerdings bräuchte man dann neue MitarbeiterInnen, die diese Aufgabe übernehmen können. Es gibt noch keine konkreten Pläne, diese Entwicklung umzusetzen.

Interessant ist hierbei besonders die Arbeitsweise mit verteilten Aufgaben im Team. Auch dass die Spracherkennungssoftware im Falle der in-house Untertitelung bei NHK kein zugekauftes Produkt ist, sondern eine Eigenproduktion, sticht heraus. Des Weiteren fällt die im Vergleich relativ hohe Untertitelungsrate und ein früher Beginn der Live-Untertitelung mit Respeaking Anfang der 2000er-Jahre auf. Wie die Analyse zeigt, bestehen in der Arbeitsweise zwischen NHK und Faith signifikante Unterschiede, sowohl in der Anzahl der beteiligten Personen als auch in der durchschnittlichen Dauer der Einsatzzeit am Stück.

6.1.2 Fähigkeiten und Hintergründe der RespeakerInnen und EditorInnen

In diesem Unterkapitel werden als für UntertitlerInnen wichtige Fähigkeiten und gewünschte beziehungsweise reale Hintergründe besprochen. Hierbei sollte auch über die Aufnahme neuer MitarbeiterInnen gefragt werden, leider konnte diese Information aber nur von einer der beiden Stellen bezogen werden. Die Ausbildung, bei der die Fähigkeiten weiterentwickelt werden, ist ebenfalls Thema.

Bei **NHK** sind viele der RespeakerInnen und EditorInnen seit vielen Jahren in dieser Position beschäftigt. Es werden eher selten neue MitarbeiterInnen eingestellt. Dies ist besonders darin begründet, dass sich diese MitarbeiterInnen über die Zeit sehr stark auf ihre Bereiche spezialisieren und einarbeiten und daher sehr schwierig zu ersetzen sind. Die Wörterbücher, die sie über die Zeit erstellt haben, sind auch sehr umfangreich und für neue Personen nicht einfach zu handhaben. Ein Beispiel hierfür sind die Sumo-Übertragungen, die bereits seit 2003 regelmäßig untertitelt werden und bei denen viele der darauf spezialisierten MitarbeiterInnen bereits seit diesem Zeitpunkt tätig sind.

Über den Interviewprozess bei der Aufnahme neuer MitarbeiterInnen bei NHK konnte ich leider keine genaue Auskunft bekommen, da dieser an eine Subfirma ausgelagert ist. Wichtige Eigenschaften für einen oder eine RespeakerIn sind laut meinen Interviewpartnern gute Sprachfähigkeiten, die Fähigkeit, Sätze im Kopf zu formulieren, und eine deutliche Aussprache. Es wird als intuitive Fähigkeit der RespeakerInnen betrachtet, gleichzeitig zuhören und sprechen zu können. Wer dies nicht kann, sei für den Beruf nicht geeignet. Auch braucht man die Fähigkeit, Sätze schnell umformulieren zu können und jene Stellen zu eliminieren, die verwirrend sein könnten, ohne die Grundbedeutung des Satzes zu verlieren. Da die Aufgaben im Gegensatz zur Untertitelung in anderen Ländern auf die RespeakerInnen und die EditorInnen

aufgeteilt sind, sind auch die Kompetenzen verteilt. Für die EditorInnen sind Rechtschreib- und Schriftzeichenkenntnisse relevant. Damit sind sehr hohe Sprachfähigkeiten und ein schnelles Denkvermögen essenziell. Da es während einer Live-Sendung normalerweise keine Möglichkeit gibt, etwas nachzuschlagen, müssen die EditorInnen alle diese Fakten im Kopf haben.

Für die Arbeit bei **Faith** wird keine besondere Berufsausbildung verlangt; ein Universitätsabschluss ist generell nicht notwendig. Vor der Aufnahme in die Firma wird ein Test durchgeführt, dieser zielt aber eher auf Allgemeinwissen ab und behandelt keine speziellen die Untertitelung betreffenden Fähigkeiten. Der wichtigste Teil des Aufnahmeprozesses ist das persönliche Interview. Die wichtigste Qualifikation für eine/n gute/n RespeakerIn ist eine gute Aussprache.

簡単な基礎的なテストをします。例えば、最近の一般常識みたいなテストはします。ただ、ほとんどが面接ですね。面接で活舌がいいとか、あの、声が出せるかどうかというので選んでます。男性、女性も関係ないです。

Wir machen einen einfachen Grundlagentest, dabei geht es um Allgemeinwissen. Aber das wichtigste ist eigentlich das Interview. Dort können wir feststellen, ob die Aussprache gut ist, und ob die Stimme klar klingt. Danach wählen wir dann aus. Es spielt auch keine Rolle, ob es ein Mann oder eine Frau ist. (Interview Faith 1: 171-173)²

Für EditorInnen ist auch die Anschlagszahl beim Schreiben wichtig. Mein Interviewpartner meinte weiter, dass jüngere Leute meistens besser für die Arbeit geeignet seien und sich schneller in diese einarbeiten würden. Es sei aber nicht notwendig, einen bestimmten Hintergrund zu haben, um diese Arbeit ausüben zu können. Wichtig sei allerdings ein Interesse an den Inhalten und Fernsehen im Allgemeinen. Dies wird im Interview besprochen, wodurch auch ermittelt werden soll, wie selbstständig eine Person beim Wissenserwerb vorgeht. Obwohl kein Universitäts- oder ähnlicher Abschluss notwendig ist, wirbt die Firma Faith gezielt künftige MitarbeiterInnen an Medienzweigen von Fachhochschulen oder Universitäten an. Mein Interviewpartner, der selbst als Respeaker arbeitet, hat ebenfalls eine Fachhochschule im

² Alle Zitatübersetzungen erfolgten durch die Autorin der Arbeit

Medienbereich absolviert, auch wenn zum Zeitpunkt seines Eintritts in die Firma dort noch keine Live-Untertitelung stattfand.

Obwohl auch andere berufliche Hintergründe vorkommen, haben die RespeakerInnen bei **NHK** fast alle zuvor als NachrichtensprecherInnen oder ReporterInnen gearbeitet, aber auch andere Hintergründe im Bereich Journalismus sind häufig und können hilfreich sein. Sie haben daher schon Erfahrung mit dem Fernsehen und dem professionellen Sprechen. Einige haben sogar in demselben Bereich gearbeitet, für den sie jetzt als RespeakerInnen tätig sind, zum Beispiel waren manche der RespeakerInnen bei Sumo-Wettkämpfen zuvor KommentatorInnen für diese Sportart. Da diese MitarbeiterInnen aber relativ hohe Personalkosten verursachen, wird verstärkt versucht, auch andere Personen einzuschulen oder die EditorInnen als RespeakerInnen einzusetzen. Eine standardisierte Einschulung gibt es in diesem Sinne nicht: die neueren MitarbeiterInnen lernen von den Erfahreneren und üben das Respeaking, bis sie für die Sendungen qualifiziert sind. Von Seiten der Firma Faith, welche ja auch für andere Fernsehsender tätig ist, wurde noch erwähnt, dass bei manchen Fernsehsendern die angehenden NachrichtensprecherInnen, bevor sie als solche tätig werden, erst als RespeakerInnen arbeiten.

Auch bei **Faith** ist die Ausbildung sehr praktisch orientiert. Das Respeaking wird mit dem amivoice-Programm geübt. Dadurch wird nicht nur sichergestellt, dass die betreffende Person mit dem Programm umgehen kann, sondern dieses verfügt auch wie bei NHK über Sprachprofile, die auf ihre NutzerInnen trainiert werden müssen. Beim selbstständigen Üben werden Sendungen angesehen und dazu gesprochen, wie es auch beim tatsächlichen Respeaking stattfindet. Es ist hierbei wichtig, die verschiedenen Genres kennen zu lernen, da man auf diese Art die notwendigen Vokabeln und Eigenheiten der Sendungen kennen lernt. Mein Interviewpartner war der erste Respeaker in seiner Firma, dementsprechend war seine Ausbildung zu Beginn noch eher wenig ausgereift. Hilfreich sei gewesen, dass er von den Fernsehsendern Informationen zu ihrem Programm erhielt. Eine gute Allgemeinbildung besonders in den Themenbereichen Wirtschaft und Politik sei sehr wichtig. Dies werde heute auch in die Ausbildung mit einbezogen, wobei das selbstständige Arbeiten und Erwerben von Wissen immer noch sehr wichtig sei.

Bei der Einschulung in die Tätigkeit als EditorIn sind vor allem die Kenntnisse der japanischen Sprache, insbesondere bezüglich der richtigen Wahl der Schriftzeichen für Eigennamen und Ähnliches, sehr relevant. Es wird darüber hinaus auch der Umgang mit einem Tastatur und das schnelle Schreiben trainiert, da die EditorInnen bei Faith mit einem konventionellen Tastatur arbeiten. Die Einschulung dauert im Durchschnitt etwa drei Monate

für RespeakerInnen und zwei Monate für EditorInnen, wobei die genaue Dauer je nach Person variiert.

Die Arbeitszeiten der RespeakerInnen variieren stark, die meisten arbeiten allerdings nicht in einer Vollzeitbeschäftigung. Vor dem Respeaking übernehmen sie auch Aufgaben der Vorbereitung. Darunter fällt das Einlesen von Begriffen in die Wörterbücher der Software, bei **NHK** aber auch die Anwesenheit bei Proben, die der Sender vor manchen Sendungen durchführt. Für die Vorbereitung erhalten sie Unterlagen zu den jeweiligen Sendungen. Der Regisseur koordiniert hierbei, wie die Vorbereitungen ablaufen und erstellt auch Listen mit benötigten Wörtern. Die RespeakerInnen können auch Hinweise darauf geben, bei welchen Wörtern Schwierigkeiten aufgetreten sind, allerdings ist dies eher die Aufgabe der EditorInnen, da diese auch häufiger während der Korrektur mit diesen Fehlern konfrontiert sind. Darüber hinaus haben die meisten RespeakerInnen keine andere Beschäftigung bei NHK, obwohl es einige wenige gibt, die weiterhin als NachrichtensprecherInnen arbeiten.

Bei **Faith** gibt es ebenfalls verschiedene Anstellungsverhältnisse, inklusive MitarbeiterInnen, die Vollzeit beschäftigt sind. Da Faith auch vorgefertigte Untertitel erstellt, sind einige auch in diesem Bereich tätig. Direkt mit dem Respeaking werden etwa drei Stunden am Tag verbracht, ansonsten befassen sich die MitarbeiterInnen natürlich auch mit der Vorbereitung auf die jeweilige Sendung und trainieren das Sprachprogramm. Interessant ist zu bemerken, dass die MitarbeiterInnen bei Faith normalerweise in beiden Positionen arbeiten, also sowohl als EditorInnen und als RespeakerInnen. Allerdings werden sie nicht während einer Sendung in beiden Rollen eingesetzt.

Meine InterviewpartnerInnen sahen Gemeinsamkeiten zwischen dem Respeaking und dem Simultandolmetschen. Besonders die Fähigkeit, gleichzeitig hören und sprechen zu können, sei für beides erheblich wichtig. Außerdem hat es Überlegungen gegeben, ausgebildete SimultandolmetscherInnen für die bis jetzt nicht verwirklichte interlinguale Untertitelung bei Faith einzusetzen.

Zusammengefasst sticht während der Interviews heraus, dass viele der MitarbeiterInnen schon lange Zeit für die Untertitelungsabteilung tätig sind, sowie die Tatsache, dass diese häufig einen Hintergrund als NachrichtensprecherInnen haben, und teilweise sogar in denselben Sendungsgenres gearbeitet haben. Spezifisch für die Untertitelung in Japan ist außerdem die Verteilung der Kompetenzen auf RespeakerInnen und EditorInnen. Während beide Sprachkompetenzen aufbringen müssen, ist bei den RespeakerInnen die Aussprache wichtig und bei den EditorInnen eher die Rechtschreibung und die Anschlagsgeschwindigkeit. Bei der Firma Faith müssen die MitarbeiterInnen natürlich trotzdem alle Kompetenzen in sich vereinen,

da die Rollen auch getauscht werden können. Allgemein wird von den MitarbeiterInnen eine hohe Selbstständigkeit in der Arbeitsweise erwartet. Interessant ist aber auch der Aufgabenbereich des Regisseurs, durch den einige Vorbereitungsaufgaben, wie das Erstellen von Wörterlisten, nicht in den Bereich der RespeakerInnen oder EditorInnen fallen.

6.1.3 Qualitäts- und formale Kriterien

In diesem Teil der Befragung sollte herausgefunden werden, ob es staatlich vorgegebene oder auch interne Qualitätsrichtlinien gibt und wie die Qualitätskontrolle generell umgesetzt wird. Auch formale Kriterien wie die Länge der Untertitel und die Standzeiten werden besprochen. Es bestehen hier ebenfalls leichte Unterschiede zwischen den verschiedenen Fernsehsendern, wozu Faith, das auch für andere Sender tätig ist, ebenfalls Auskunft geben konnte.

Es werden momentan weder bei NHK noch bei Faith besondere Formeln zur Qualitätskontrolle angewendet und es gibt auch keine einheitlichen Richtlinien hierzu. Vom Staat werden ebenfalls keine Mindestanforderungen bezüglich der Qualität vorgegeben. Eine ständige Qualitätskontrolle findet aber bei jeder Sendung durch eine/n MitarbeiterIn statt, der/die die Sendung mithört und bei einer anschließenden Besprechung auf Fehler hinweist. Es handelt sich also eher um eine qualitative als eine quantitative Herangehensweise. Es gibt laut Aussagen der Interviewpartner kaum negative Rückmeldungen der Zielgruppen. Sollten die Untertitel mehr als 10 Minuten ganz ausfallen, gilt dies laut dem Regisseur von NHK gesetzlich als ein „Sendungszwischenfall“ und es muss eine Entschuldigung seitens des Senders und eine erneute Ausstrahlung der Sendung erfolgen.

Die Verzögerung der Untertitel kann stark variieren. Ist die Erkennungsrate sehr schlecht, kann eine Verzögerung von bis zu 30 Sekunden entstehen. Dies ist aber nur sehr selten der Fall, durchschnittlich beträgt die Verzögerung bei **NHK** nur etwa 6 Sekunden. Es gibt keine Vorschrift für die maximale Verzögerung der Untertitel, auch wenn diese so gering wie möglich gehalten werden soll. Wichtig ist natürlich, dass Bild und Text zusammenpassen. Hier müssen die RespeakerInnen im Bedarfsfall Kürzungen vornehmen. Mein Gesprächspartner bei **Faith** gab die durchschnittliche Verzögerung mit 10 Sekunden an. Bei Faith werden regelmäßig Tests dazu durchgeführt, welche Erkennungsrate die einzelnen MitarbeiterInnen erreichen; es gibt auch Richtwerte dazu.

Ehe ein/e RespeakerIn bei **Faith** die erste Live-Sendung durchführen darf, muss er oder sie bei einem Test eine ausreichende Erkennungsrate erreichen. Dazu wird eine Nachrichtensendung mit vorhandenem, korrektem Drehbuch herangezogen und mit den

produzierten Untertiteln verglichen. Allerdings ist die zu erreichende Schwelle nicht ganz genau festgelegt, sondern die Beurteilung hängt von den erfahreneren MitarbeiterInnen ab. Das Kriterium ist hierbei eher, ob der fertige Text von den EditorInnen in absehbarer Zeit korrigiert werden könnte oder ob vermutlich eine zu hohe Verzögerung entstehen würde. Auch bei bereits etablierten RespeakerInnen werden ab und zu solche Tests durchgeführt, um ihre Leistungen laufend bewerten zu können. Allerdings ist dies bei langedienten RespeakerInnen nicht mehr so häufig der Fall wie bei den neu hinzugekommenen. Es werden außerdem Vergleiche mit anderen Untertitelungsfirmen durchgeführt, indem eine von diesen untertitelte Sendung bei Faith versuchsweise erneut untertitelt wird, um dann die Ergebnisse zu vergleichen.

Mit einer absichtlichen Verzögerung zwischen Aufzeichnung und Ausstrahlung von Live-Sendungen, um den UntertitlerInnen mehr Zeit zu gewähren, wurde nur einmal bei **NHK** experimentiert, und es wurde in diesem Fall auch nicht die Respeaking-Methode, sondern die Tastatur-Methode angewandt. Die absichtliche Verzögerung betrug hierbei 30 Sekunden. Dies hat für die Tastatur-Methode gereicht, allerdings müsste bei der Respeaking-Methode vermutlich eine noch längere Verzögerung eingeführt werden. Jedenfalls konnte sich diese Herangehensweise nicht durchsetzen, auch deshalb, weil sie sehr viel Koordination zwischen den verschiedenen Abteilungen voraussetzt und damit eine logistische Herausforderung darstellt.

Bezüglich der formalen Kriterien gibt es folgende Vorgaben: Die Länge der Untertitel beträgt durchschnittlich 15-16 Zeichen pro Zeile und es können bis zu zwei Zeilen gesendet werden. Eine Untertitelzeile soll dabei mindestens 3 Sekunden auf dem Bildschirm bleiben.

Sprecherwechsel werden mit dem Namen und dem „<<“ Symbol angezeigt. Für die Identifizierung bestimmter SprecherInnenrollen werden Farben verwendet, zum Beispiel, um zwischen ModeratorIn und Gast zu unterscheiden. Der/die RespeakerIn aktiviert diese, indem er/sie während des Sprechens einen Knopf drückt. Satzzeichen fügt das Programm normalerweise automatisch ein, der/die EditorIn kann allerdings nach eigenem Ermessen auch noch zusätzliche Zeichen einfügen. Das Einfügen von Satzzeichen hängt auch mit dem Zeichenlimit zusammen, das nicht überschritten, aber auch gut ausgenutzt werden soll. Auch hiernach wird entschieden, ob ein Einfügen von Satzzeichen sinnvoll ist oder eher nicht.

Die Wörtlichkeit der Wiedergabe wird als wichtiges Kriterium angeführt, allerdings muss sie hinter der Verzögerung zurückstecken – wenn diese zu hoch wird, werden Kürzungen durchgeführt. Auch werden unwichtige Füllwörter häufig von den RespeakerInnen entfernt. Es ist bekannt, dass eine große Heterogenität in den Erwartungen der Zielgruppe vorherrscht. Einige haben kein Problem mit Verkürzungen in der Untertitelung, wenn dadurch Fehler und

die Verzögerung vermindert werden. Andere jedoch wünschen sich in jeden Fall eine hohe Wörtlichkeit, auch wenn die Lesbarkeit der Untertitel darunter leidet.

Bei **Faith** wurde die Wörtlichkeit der Untertitel als stark variabel beschrieben. Laut dem persönlichen Ermessen des einen Interviewpartners liegt sie durchschnittlich aber bei 90 bis 95 Prozent, was ebenfalls eine Spanne aufzeigt. Abweichungen passieren vor allem dann, wenn mehrere Leute zugleich sprechen oder der/die SprecherIn sich wiederholt. Wenn Personen erst während des Sprechens ihre Sätze formulieren, kommen viele Wiederholungen vor und es liegt im Ermessen des/der RespeakerIn, diese neu zu formulieren. Die Korrektheit hat generell Vorrang im Gegensatz zur Wörtlichkeit und wird auch als wichtiger erachtet als die Verzögerung. Es kommt hierbei aber auch auf die Art der Sendung an. Wie bereits erwähnt, ist die Wörtlichkeit von Nachrichtensendungen besonders wichtig, während bei anderen Programmen teilweise gekürzt wird.

Faith erhält kein direktes Feedback von den Zielgruppen, erfährt aber von jenem Feedback, das die VertreterInnen von NHK bei Konferenzen bekommen. Auch hier ist man sich der Heterogenität der Anforderungen bewusst. Manche Mitglieder der Zielgruppen würden sich beispielsweise wünschen, dass auch Füllwörter in den Untertiteln wiedergegeben werden. Farben zur Kennzeichnung einzelner Personen scheinen nicht besonders beliebt zu sein, die ZuseherInnen wünschen sich eher eine Kennzeichnung durch Namen. Bezüglich der Geschwindigkeit gäbe es keine Beschwerden, früher sei dies allerdings schon vorgekommen. Interessant zu bemerken ist, dass mein Interviewpartner angab, die Untertitel bei NHK seien generell langsamer als bei anderen Fernsehsendern. Bei Privatsendern müsse ein Untertitel nur 2 Sekunden stehenbleiben und nicht die bei NHK gewünschten 3 Sekunden. Dies ist auch deshalb bei NHK besser realisierbar als bei den Privatsendern, weil die SprecherInnen zum Beispiel bei den Nachrichten striktere Vorgaben haben, in welchem Tempo sie sprechen müssen. Wenn aber bei den Privatsendern die Personen im Fernsehen schneller sprechen, müssen zwangsläufig auch die Untertitel schneller sein, um mithalten zu können. Es wurde die Meinung geäußert, dass die Geschwindigkeit kein großes Problem für die Zielpersonen sei, da diese generell eine hohe Lesegeschwindigkeit hätten.

Anders als in manchen anderen Ländern gibt es zwar Vorgaben zur Untertitelungsmenge, aber nicht zu Qualität der Untertitelung. Diese wird eher durch qualitative Methoden überprüft und orientiert sich darüber hinaus am Feedback der ZuseherInnen. Die Verzögerung von sechs Sekunden liegt etwa im Durchschnitt der in der Literatur besprochenen Verzögerungen. Interessant bei den formalen Kriterien ist, dass es dahingehend Unterschiede zwischen öffentlichen und Privatsendern gibt, welche Standzeiten die Untertitel haben müssen.

Generell wird bei den Untertiteln eine hohe Wörtlichkeit bevorzugt, obwohl diese trotzdem variieren kann. Das Feedback der Zielgruppen ist, wie so häufig, eher gemischt, worauf auch das Kapitel zur Zielgruppenbefragung noch eingehen wird.

6.1.4 Zukunft der Untertitelung in Japan

In diesem Unterkapitel wird eine Einschätzung der Interviewpartner zur Weiterentwicklung und den Zukunftsaussichten der Live-Untertitelung beschrieben. Hierbei ist natürlich zu beachten, dass die von mir befragten Personen nicht unbedingt die Hauptentscheidungsträger in diesem Zusammenhang sind und daher über vieles nur Vermutungen abgeben konnten.

Die Pläne für die Zukunft von **NHK** beinhalten eine Ausweitung der Untertitelung. Momentan sind Untertitel nur zwischen 7 Uhr und 24 Uhr verfügbar. Auch innerhalb der Zeiten sollen Untertitel korrekter und weniger verzögert werden. Andere Formate, wie zum Beispiel interaktives Fernsehen, miteinzubeziehen, ist ebenso ein Ziel.

Japanisch wird als eine sehr schwierig zu untertitelnde Sprache erachtet, besonders mit den Methoden der Spracherkennung. Daher werden die bisherigen Erfolge von NHK sehr hoch bewertet. Trotz der auftretenden Schwierigkeiten halten meine Interviewpartner die Spracherkennung für die Zukunft der Untertitelung. Durch bessere Sprachmodelle und Wörterbücher soll die Erkennungsrate auf durchgehend 95 Prozent angehoben werden. Wenn diese Rate erreicht werden würde, könnte die Spracherkennung die Tastaturmethode als zuverlässigste Methode verdrängen. Abgesehen davon soll die Methode auf die interlinguale Untertitelung ausgedehnt werden.

Allerdings muss die Verwendung von Spracherkennung aber nicht die Verwendung von Respeaking voraussetzen. Es wird schließlich auch zum jetzigen Zeitpunkt schon direkte Spracherkennung verwendet, zum Beispiel für die Texte von NachrichtensprecherInnen. Dieser Text wird dann allerdings noch von zwei EditorInnen ausgebessert. Obwohl die automatische Erkennung in Zukunft häufiger genutzt werden soll, waren meine Interviewpartner nicht der Meinung, dass Untertitelung je komplett ohne menschliche Intervention funktionieren wird. Auch eine reine Verwendung der direkten Erkennung mit Verbesserung der EditorInnen würde wohl noch eine Weile dauern:

ダイレクトになったものこう力で修正いくっていうこの校正者のほうが重要になってくるんじゃないかな。ただ日本語は難しいので中々ダイレクトというのは非常にまだまだ時間かかるんじゃないかな。先言った通りニュースは一人のアナウンサーがゆっくりしゃべるからダイ

レクトにできるんですが、複数の話者が一斉に話始めるとダイレクト認識は非常に難しいですね。

Wenn wir auf direkte Spracherkennung umstellen würden, wären dann eher die EditorInnen wichtig [und nicht die RespeakerInnen, Anm.]. Aber weil Japanisch sehr schwierig ist, dauert es ziemlich lange, bis man die direkte Methode verwenden kann. Wie ich vorher schon sagte, kann diese Methode nur für die Texte von NachrichtensprecherInnen verwendet werden, wenn diese alleine sprechen. Sobald mehrere Sprecher in einem Gespräch gleichzeitig sprechen, wird die Spracherkennung sofort schwieriger. (Interview NHK 632-636)

Bei **Faith** sieht man die Zukunft der Untertitelung ebenfalls in der Spracherkennung. Hier wird momentan darüber gesprochen, inwieweit smartphonebasierte Programme für die Spracherkennung bei der Untertitelung eingesetzt werden können. Dies wird aber momentan aufgrund von Unsicherheiten nicht wirklich praktiziert. Die Genauigkeit der Spracherkennungsprogramme werde in Zukunft wohl viel besser werden, eine Notwendigkeit von Menschen im Prozess sei aber weiterhin gegeben, insbesondere deshalb, weil Fehler, die durch eine Maschine entstehen, von Menschen häufig weniger nachvollzogen werden können als jene, die ein Mensch machen würde, wodurch tendenziell das Verständnis stärker beeinträchtigt wird. Obwohl die geschlossene Meinung herrscht, dass Menschen im Prozess in nächster Zeit nicht durch künstliche Intelligenz zu ersetzen sind, gibt es doch bezüglich der Sicherheit ihrer Jobs eine gewisse Unsicherheit unter den RespeakerInnen.

Zusammengefasst soll in Zukunft hauptsächlich versucht werden, bevorzugt die automatische Spracherkennung einzusetzen, um Kosten zu sparen. Dies stellt eine Unsicherheit für die RespeakerInnen bezüglich ihres Arbeitsplatzes dar. Dennoch werden weiterhin vermutlich MitarbeiterInnen notwendig sein. Darüber hinaus soll die Untertitelung auf andere Programme und andere Uhrzeiten ausgeweitet werden. Dazu gibt es auch neue Richtlinien der Regierung, die in den oberen Kapiteln bereits besprochen worden sind. Weiters gibt es Überlegungen, die verwendete Technik nach Möglichkeit zu vereinfachen.

6.1.5 Schwierigkeiten und Lösungsansätze

Besonders interessant im Vergleich zur Untertitelung in anderen Ländern sind die auftretenden Probleme und Lösungsansätze. Besonders hier spielen die Eigenheiten der japanischen Sprache eine große Rolle.

Am häufigsten treten Schwierigkeiten wegen der im Japanischen sehr verbreiteten Homonyme auf. Dies wurde von den Vertretern von **NHK** als größte Fehlerquelle genannt. Diese betreffen meistens die EditorInnen, da sie die Homonyme korrekt auswählen müssen. Technische Probleme kommen hingegen weniger häufig vor. Dies wird damit begründet, dass schon seit so langer Zeit dasselbe Programm verwendet wird und dadurch die meisten Fehler bereits ausgebügelt sind. Wenn es zu technischen Problemen kommt, dann treten diese meistens mit den Touchscreens der EditorInnen auf.

Eine stressige Situation für die MitarbeiterInnen sind Sendungen, für die es wenig Vormaterial gibt oder in welchen plötzlich eine Live-Schaltung erfolgt. Manchmal wissen die MitarbeiterInnen zwar, dass und wohin es eine Live-Schaltung geben wird, erhalten aber keine Informationen dazu, was dort besprochen werden soll. In diesem Fall recherchieren die RespeakerInnen und EditorInnen noch kurz vorher zu dem Ort und versuchen herauszufinden, was Inhalt der Sendung sein könnte, um Fehler zu vermeiden.

Wie aufgrund der Sensibilität des Themas zu erwarten gewesen war, fiel es meinen Interviewpartnern nicht leicht, Fragen nach Schwierigkeiten zu beantworten. Es ist allerdings auch vorstellbar, dass dies teilweise auch damit zusammenhängt, dass die erfahreneren MitarbeiterInnen viele der Probleme nicht mehr als so schwerwiegend wahrnehmen. Eine schwierige Situation, die von meiner Kontaktperson bei **Faith** genannt wurde, ist das Untertiteln während Katastrophensituationen. Wenn etwa ein Erdbeben oder Taifun auftritt, wird ohne lange Vorwarnung auf eine Live-Sendung umgeschaltet und die RespeakerInnen haben kaum Zeit, sich darauf vorzubereiten. Dazu kommt, dass diese Sendungen dann häufig eine sehr lange Zeitdauer aufweisen. Diese werden auch nicht durch eine Werbepause unterbrochen, welche die RespeakerInnen als Pause nutzen könnten. Hierbei kann es sogar vorkommen, dass die MitarbeiterInnen länger als drei Stunden durchgehend arbeiten müssen. Obwohl die RespeakerInnen und EditorInnen in diesem Zeitraum nach Möglichkeit abgelöst werden, haben sie häufig lange Zeit keine Möglichkeit, auch nur eine kurze Pause zu machen.

Es gibt auch keine vorgegebenen Intervalle, in denen ein Wechsel stattfindet, sondern es hängt davon ab, wie sich der/die RespeakerIn fühlt. Wenn es keine Werbeunterbrechungen gibt, werden diese Pausen an den Aufbau der Sendung angepasst, um zu möglichst wenigen Unterbrechungen zu führen. Dies ist kann dann sehr anstrengend für die MitarbeiterInnen werden:

飲み物も飲まなかったり。あるいは、おなかがすいちゃったり、トイレに行けなかったり。そういうやっぱり急なトラブルみたいなものにそういう放送しなきゃいけないのでわれわれね。そういう時が一番大変かもしれないですね。

Wenn es plötzlich eine Notfallsituation gibt, dann müssen wir diese Art Sendung machen. Wir können dann nichts trinken und nichts essen und auch nicht zur Toilette gehen. Das sind wohl die schwierigsten Situationen. (Interview Faith 2 174-176)

Es sei auch manchmal schwierig, da man nicht das Ende des Satzes abwarten kann, bevor man zu sprechen beginnt. Wenn man sich zu viel auf den Inhalt konzentriert, können Verzögerungen entstehen. Zugleich müsse man aber natürlich auch die Sinnhaftigkeit im Auge behalten. Eine weitere schwierige Situation ist, wenn Probleme mit der Technik beziehungsweise mit dem Computer auftreten, besonders wenn diese nur von einem Außenstehenden gelöst werden können.

Eine Strategie zur Vermeidung von Fehlern ist die richtige Vorbereitung. Um sicherzustellen, dass die Umwandlung der Laute in Schriftzeichen korrekt erfolgt, werden Eigennamen oder Wörter, die voraussichtlich in der Sendung auftauchen werden, in das System eingelesen. Hierbei merkt sich die Software auch, wie häufig diese vorkommen, und wird häufig verwendete Namen dann eher den EditorInnen anbieten. Sollten Namen vorkommen, die nicht vorbereitet wurden und bei denen die Schreibweise nicht klar ersichtlich ist, werden diese in der Silbenschrift dargestellt. Die ausführliche Vorbereitung hilft außerdem bei der Bewältigung von Stress während der Arbeit.

In problematischen Situationen kommt den RespeakerInnen und EditorInnen aber auch eine Eigenheit des japanischen Fernsehens zugute. Während Interviews und in Talkshows werden die SprecherInnen häufig schon von vornherein mit *open captions* untertitelt, die für jeden sichtbar sind. Wenn also die Live-Untertitelung schwieriger wird oder Fehler entstehen, kann der Inhalt häufig durch die *open captions* verstanden werden. In manchen Fällen ist eine weitere Untertitelung sogar gar nicht notwendig.

Die beiden Hauptschwierigkeiten stellen somit die häufig auftretenden Homonyme und die Liveschaltungen in Katastrophensituationen dar. Wichtig um Fehler zu vermeiden, ist besonders die gute Vorbereitung und das Einlesen eventuell problematischer Worte in die Software.

6.1.6 Fazit und Vergleich

In vieler Hinsicht ähneln die Aussagen meiner Interviewpartner jenen aus zum Thema Respeaking in anderen Ländern veröffentlichten Arbeiten. Ihnen zufolge hat sich Japan noch nicht für eine einheitliche Methode der Live-Untertitelung entschieden. Momentan ist die Tastatur-Methode noch vorherrschend, doch die Spracherkennung gewinnt immer mehr an Bedeutung und wird trotz auftretender Schwierigkeiten als zukunftssträchtige Methode betrachtet. Auffällig ist, dass die RespeakerInnen nicht alleine, sondern gemeinsam mit den EditorInnen arbeiten. Dies ist im europäischen Raum in den meisten Ländern nicht der Fall (Romero-Fresco 2011) und stellt daher einen Unterschied in der Arbeitsweise dar. Bei NHK gibt es eine hohe Untertitelungsrate durch den Versuch, die Richtlinien der japanischen Regierung umzusetzen. Diese betreffen aber nur eine bestimmte Zeitspanne zwischen 7 Uhr morgens und 12 Uhr nachts. Auch wurde mit der Live-Untertitelung zu einem ähnlich frühen Zeitpunkt begonnen, wie es auch bei den in Europa als Vorreiter geltenden Ländern Großbritannien und den Niederlanden der Fall war.

Die Qualifikationen der RespeakerInnen decken sich auch in einigen Punkten mit den in der Literatur erwähnten, wie zum Beispiel in Bezug auf die sprachliche Kompetenz und die technischen Fähigkeiten sowie das gleichzeitige Zuhören und Sprechen. Ein Unterschied liegt zumindest bei den NHK in-house Untertiteln in der Rollenaufteilung, durch die auch die Kompetenzen aufgeteilt sind. Wie in den meisten in der Literatur beschriebenen Ländern findet die Ausbildung für die Live-Untertitelung ausschließlich bei dem jeweiligen Fernsehsender beziehungsweise in der Untertitelungsfirma statt. Die Hintergründe der MitarbeiterInnen unterscheiden sich stark, obwohl viele aus anderen Abteilungen beim Fernsehen oder aus Ausbildungen im Medienbereich kommen. Es gibt also wie in der Literatur zu den Kompetenzen beschriebenen (siehe Kapitel 3.1) bei NHK und Faith ebenfalls keine bestimmte Ausbildung, die für die Arbeit als RespeakerIn als Voraussetzung gesehen wird. Parallelen zwischen Simultandolmetschen und Respeaking werden durchaus gesehen, da diese aber nur als Antwort auf eine in diese Richtung formulierte Frage ausgesprochen und nicht spontan geäußert wurden, ist fraglich, wie stark diese Verbindung wirklich präsent ist.

Es gibt keine strikten Vorgaben bezüglich der Qualität, und es werden auch keine Formeln benutzt, um diese zu berechnen. So war auch das NER-Model nach Romero-Fresco den befragten Personen als Berechnungsmethode unbekannt. Dennoch finden Qualitätskontrollen statt und es wird im Ermessen der kontrollierenden Personen auch eine Konsequenz aus einer schlechten Erkennungsrate gezogen. Die formalen Vorgaben bezüglich

Länge und Stehzeiten der Untertitel scheinen zu variieren. Grundsätzlich ist die Länge aufgrund des stark unterschiedlichen Schriftsystems ohnehin schwer mit europäischen Sprachen vergleichen. Es gab in der letzten Zeit kaum Beschwerden über zu schnelle Untertitel und abgesehen von der Verzögerung scheint die Geschwindigkeit nicht als großes Problem gesehen zu werden. Zur Identifikation von Personen werden vor allem Namen verwendet, manchmal werden allerdings auch Farben genutzt.

Obwohl Schwierigkeiten mit Homonymen auch in anderen Sprachen auftreten können, sind diese im Japanischen besonders relevant. Die Eigenheit der japanischen Sprache besonders viele Homonyme aufzuweisen, führt zu besonderen Herausforderungen bei der Live-Untertitelung durch Respeaking, da diese durch die Spracherkennung nicht automatisch unterschieden werden können. Daher wurden sie von den Befragten auch besonders als Problemquelle hervorgehoben. Andere Probleme, wie zum Beispiel in Bezug auf Live-Sendungen mit geringer Vorinformation oder die Wichtigkeit der Vorbereitung, decken sich aber mit jenen Herausforderungen, die in der Literatur beschrieben wurden.

6.2 Zielgruppenbefragung

Wie bereits erwähnt wurde die Zielgruppenbefragung bei zwei verschiedenen Gehörlosen- und Schwerhörigenorganisationen durchgeführt. Es konnten auch von beiden Stellen mehrere Personen für die Beantwortung gefunden werden. Durch die Art der Befragung, die in diesem Fall über eine vermittelnde Person ablief, wurden einige der Fragen aus dem Originalfragebogen leicht verändert beziehungsweise erweitert. Da allerdings keine Vergleichbarkeit im quantitativen Sinne angestrebt wurde, ist diese Situation, wenn auch vielleicht nicht ideal, in Bezug auf die Ergebnisse der Forschung tolerabel. Obwohl die Antworten auf die von der Kontaktperson hinzugefügten Fragen in manchen Fällen interessant waren, entsprechen sie nicht wirklich dem Ziel der Forschung und werden daher in dieser Beschreibung nicht beachtet. Wo die Fragen im Vergleich zum ursprünglichen Fragenbogen verändert wurden, werden die Unterschiede im Folgenden explizit erwähnt. Es wurde auch einige zusätzliche Fragen hinzugefügt, die als für diese Forschung nicht relevant betrachtet wurden. Sie werden daher hier nicht besprochen, scheinen aber der Vollständigkeit halber in den Anhängen mit auf.

6.2.1 Demographische Daten

Insgesamt beantworteten acht Personen meinen Fragebogen, drei aus der JFD und fünf aus der Zennancho-Organisation. Die Befragten der JFD waren in ihren Vierzigern, es handelte sich um zwei Frauen und einen Mann. Bei der zweiten Befragung waren die TeilnehmerInnen drei Männer und zwei Frauen. Im Alter lagen die betreffenden Personen zwischen 34 und 58 Jahren. Insgesamt war das Geschlechterverhältnis also ausgeglichen und es konnten auch Angehörige verschiedener Altersgruppen befragt werden. Bis auf eine Person wiesen alle Befragten einen hohen Grad an Gehörlosigkeit auf (90 - 100 Dezibel Hörverlust). Fast alle Personen nutzen Hörgeräte oder Cochlea-Implantate in mindestens einem Ohr. Nur eine Person verwendet seit vielen Jahren keine Hörgeräte mehr. Auch der Zeitpunkt des Gehörverlustes war bei allen TeilnehmerInnen relativ früh. Fünf TeilnehmerInnen gaben an, von Geburt an gehörlos zu sein, während bei den anderen drei TeilnehmerInnen die Gehörlosigkeit im Alter von drei bis vier Jahren begann. Im Folgenden werden die TeilnehmerInnen in der Reihenfolge der Beantwortung als T1 bis T8 bezeichnet. Die TeilnehmerInnen T1-T3 sind Mitglieder bei JFD, T4-T8 bei Zennancho.

Bezüglich der Muttersprache gaben T1 und T3 sowohl Japanisch als auch Japanische Gebärdensprache als ihre Muttersprache an. Person T2 gab an, nur Japanische Gebärdensprache als Muttersprache zu betrachten. Die fünf Personen aus der zweiten Befragung gaben als Muttersprache alle Japanisch an, und nicht Japanische Gebärdensprache. Bei der Zennancho-Befragung wurde in einer separaten Frage nach ihrer Kommunikationssprache gefragt. Die Antworten hier waren sehr weitläufig, bei den meisten wurde Lautsprache und schriftliche Kommunikation genannt. Einige nannte auch gebärdengestütztes Japanisch (dieses ist nicht deckungsgleich mit der Japanischen Gebärdensprache). Gebärdensprache als solche wurde aber im Gegensatz zu den TeilnehmerInnen aus der ersten Gruppe nicht genannt.

Während das Fernsehverhalten als solches bei den TeilnehmerInnen unterschiedlich war, gaben fast alle an, wenn sie Fernsehen würden, täten sie dies nur mit Untertiteln, und sie gaben an, dass sie keine Art von Fernsehsendungen ohne Untertitel ansehen würden. In Bezug auf die Frage, ob und für welche Sendungen sie Untertitel nutzen, gab es in der zweiten Gruppe leichte Unterschiede in der Beantwortung. Die meisten TeilnehmerInnen antworteten, in jedem Fall Untertitel zu nutzen, sowohl für Live-Sendungen als auch für jene, die voraufgezeichnet wurden. T5 hingegen gab an, diese für Live-Sendungen nicht zu nutzen, während T8 sie manchmal nutzt und zu anderen Zeiten nicht. Die meisten TeilnehmerInnen waren der Meinung, dass es keine Sendungen gibt, für die keine Untertitel notwendig sind. Nur eine Person meinte, dass es

Sendungen gibt, für die keine Untertitel notwendig sind. Diese Person war es auch, die für Live-Sendungen keine Untertitel verwendet.

6.2.2 Qualität der Untertitelung

T1 gibt an, dass die momentane Untertitelung nicht ausreichen würde. Zwar seien die meisten im Voraus aufgenommen Sendungen untertitelt, aber die Untertitelung würde bei manchen Live-Sendungen, insbesondere Variety-Sendungen, häufig fehlen. Darüber hinaus bemängelte sie fehlende Untertitel bei den Werbeunterbrechungen. Andere TeilnehmerInnen wünschten sich ebenfalls, dass in Zukunft alle Programme Untertitel haben sollten, was momentan noch nicht der Fall sei. T5 merkte an, sich eine vollständige Untertitelung aller Sendungen zu wünschen, egal durch welche Übertragungsform diese übermittelt werden. Die TeilnehmerInnen nannten auch verschiedene Sendungen, für die sie sich besonders dringend Untertitelung wünschen. Hierbei wurden vor allem jene Sendungen genannt, die spät in der Nacht ausgestrahlt werden, da ja nur bis Mitternacht Untertitelung stattfindet. Es wurde weiter bemängelt, dass zwar die Nachrichten um 7 Uhr untertitelt sind, die Frühnachrichten um 6 Uhr allerdings nicht, obwohl auch dort für die Betroffenen wichtige Inhalte ausgestrahlt werden könnten. Weiters wurden Parlamentssitzungen und Regionalsendungen als Problem erwähnt. Auch bei Katastrophenübertragungen kommt es laut T7 immer noch zu nicht untertitelten Sendungen, die ein Problem darstellen. Alle TeilnehmerInnen waren sich also einig, dass die Menge an Untertiteln generell nicht ausreicht.

Die Qualität der vorhandenen Untertitel wurde hingegen von allen TeilnehmerInnen als gut eingeschätzt, T1 gab zusätzlich an, dass es in den letzten Jahren eine Verbesserung gegeben habe, was die Qualität betrifft. T3 gab allerdings einige Verbesserungsvorschläge an:

時々、字幕が遅れる時があるので、リアルタイム[...]で理解できれば良い。最近ではテロップも増えており、テロップはリアルタイムで付加されているのでわかりやすいものが多い。

Manchmal sind die Untertitel verzögert, bei der Live-Untertitelung wäre es wirklich gut, wenn man diese besser verstehen könnte. In letzter Zeit werden häufig Texteinblendungen mit weiteren Informationen unten am Bildschirm eingeblendet, wenn man diese auch bei Live-Sendungen verwenden würde, das würde das Verständnis bei vielen Dingen verbessern. (Interview T3:27-30)

Ein Kritikpunkt, der allerdings ebenfalls geäußert wurde, betraf die Platzierung und Größe der Untertitel und dass durch diese Teile des Bildes verdeckt würden. Ein Vorschlag zur Verbesserung war, die Größe und Platzierung der Untertitel selbst wählen zu können. Eine weitere Kritik ging darauf ein, dass es keine einheitlichen Standards für die formalen Kriterien der Untertitelung gibt:

テレビ局によって字幕の見やすさが異なるため、更なる標準化を目指した方が良いのではないか

Da die Sichtbarkeit der Untertitel je nach Fernsehsender unterschiedlich ist, denke ich, dass die Richtlinien weiter standardisiert werden sollten.
(Interview T6:20)

Eine weitere Frage befasste sich mit der Qualität speziell von Live-Untertiteln, um zu erfassen, ob und welche Unterschiede festgestellt wurden. Hier waren die Antworten tatsächlich nicht so eindeutig positiv. Bemängelt wurde vor allem die Verzögerung (die T1 übrigens mit durchschnittlich 3 Sekunden angab, also etwas geringer als meine anderen Quellen). Dadurch, dass dann Bild und Ton teilweise nicht mehr zusammenpassen würden, können Verständnisprobleme entstehen, insbesondere bei Nachrichtensendungen. T2 meinte, es gäbe manchmal aufgrund von Homonymen falsch dargestellte Schriftzeichen, wodurch die Wortbedeutung verändert werden würde, dies wäre für sie aber eigentlich kein großes Problem. T5 wiederum merkte an, dass es eventuell helfen würde, die Untertitel durch den Einsatz von künstlicher Intelligenz zu erstellen.

Ein weiterer Aspekt, der aufgrund der Verzögerung problematisch sein kann, ist, dass die Untertitel aufhören und abgeschnitten werden, wenn die Sendung in die Werbepause übergeht. Hierbei fehlt dann durch die Verzögerung der letzte Teil der Untertitelung, daher wäre es wünschenswert, diese verbleibenden Untertitel trotz der Unterbrechung noch auszustrahlen. Eine Person erwähnte auch das Auftreten von Fehlern in den Untertiteln. Als Lösungsansatz wurde die absichtliche Verzögerung der Sendungen genannt. Auch bei den auftretenden Problemen wurden wieder in etwa dieselben Anliegen bezüglich Standardisierung und der Werbepause wiederholt, die eine starke Beeinträchtigung darzustellen scheinen.

生放送の字幕は、CMに入るときや場面が変わるときなど、字幕のしっぽをちょん切って終わらせてしまうので、話の内容を最後まで完全に理解することができない。

Wenn bei Live-Sendungen in die Werbepause geschaltet wird, oder sich der Bildschirm verändert, wird das Ende des Untertitels abgeschnitten, und ich

kann dem Inhalt des Gesprochenen leider nicht bis zum Ende folgen.
(Interview T7:22)

Das vorrangigste Problem bei der Live-Untertitelung scheint also die Verzögerung der Untertitel zu sein. Ein weiteres Problem, dass von T1 angegeben wurde, ist die Heterogenität der auf den Fernsehgeräten vorhandenen Funktionen. Manche Geräte könnten die Untertitel unterschiedlich platzieren (zum Beispiel außerhalb des Bildes), andere wiederum nicht. Die anderen Personen gaben an, abgesehen von dem zuvor erwähnten keine Probleme mit den Untertiteln zu haben. T2 führte dies auch darauf zurück, an die Verwendung von Untertiteln gewöhnt zu sein.

Im nächsten Teil wurden die TeilnehmerInnen befragt, was sie an der Untertitelung gerne verändern würden. Während die meisten TeilnehmerInnen ähnliche Kommentare dazu abgaben, wie sie zuvor auch erwähnt hatten, befassten sich zwei TeilnehmerInnen mit Themen, die nicht direkt mit der Untertitelung als solcher zu tun haben, sondern eher mit der Ausführung. T5 wünschte sich, dass die Untertitelsymbole, die darauf hinweisen, dass es Untertitel gibt, häufiger und konsistenter verwendet würden. Außerdem hätte sie gerne einen regeren Austausch zwischen den Zielgruppen und den UntertitelerstellerInnen. T7 wirft einen sehr interessanten Aspekt auf, der in dieser Arbeit bisher noch nicht berücksichtigt wurde, nämlich jenen der Miteinbeziehung von Menschen, die sowohl gehörlos als auch blind sind. Diese bräuchten eine andere Formatierung der Untertitel, um sie ebenfalls nutzen zu können, was momentan noch nicht gegeben ist. Von T1 wurde auch noch die Möglichkeit, zwischen gekürzten und vollständigen Untertiteln wählen zu können, als wünschenswert erwähnt. T3 führt als einziges Problem erneut die Verzögerung an. Obwohl die Qualität im Großen und Ganzen als hoch eingeschätzt wird, gibt es unterschiedliche Einschätzungen dazu, wie viele Probleme dennoch auftreten. Die Verzögerung wurde mehrmals als Problem erwähnt und scheint damit einer der wichtigsten Kritikpunkte zu sein.

Um näher auf das Thema der Qualität einzugehen, befasste sich eine weitere Frage damit, wie gute Untertitel im Allgemein aussehen beziehungsweise welche Aspekte dabei wichtig sind, da diesbezüglich die Meinungen auch auseinander gehen könnten und dementsprechend interessant sind. T1 äußerte sich bezüglich der nichtsprachlichen Elemente der Untertitel und strich in ihrer Antwort die Wichtigkeit derselben heraus.

音を文字化することで、雰囲気を伝えることが大切。音の説明を詳しくしてほしい場合があります。

Es ist wichtig die Atmosphäre zu vermitteln, indem die Geräusche in Text umgesetzt werden. Es gibt Situationen, in denen ich mir eine genaue Erklärung der Geräusche wünsche. (Interview T1:43-44)

T2 hingegen gab keine besonderen Kriterien für die Verbesserung an, sondern meinte allgemein, es könnten nur dann gute Untertitel erreicht werden, wenn eine Zusammenarbeit mit den Zielgruppen erfolgen würde. Interessanterweise empfand T3 die Bearbeitung von Untertiteln durch künstliche Intelligenz als problematisch und gab an, sich für eine gute Untertitelung weiterhin den Einsatz von menschlichen UntertitlerInnen zu wünschen.

Die Antworten variierten im Allgemeinen relativ stark. Mehrere Personen wünschten sich erneut eine bessere Einbindung in den Prozess der Untertitelung. Eine Person meinte, dass die UntertitelerstellerInnen mehr Unterstützung benötigen würden, machte aber nicht genau klar, welche Art von Unterstützung hier gemeint ist. Gute Untertitel würden vor allem durch Benutzerfreundlichkeit, Korrektheit und Austausch mit den Zielgruppen gewährleistet werden. Ebenfalls erneut erwähnt wurden eine geringe Verzögerung und hohe Genauigkeit der Untertitel.

Anschließend wurden die TeilnehmerInnen aufgefordert, die Untertitelung verschiedener Fernsehsender zu vergleichen und zu bewerten. Hier wurden von den meisten TeilnehmerInnen keine signifikanten Unterschiede angegeben. Auch T3 wollte keine Unterschiede in der Qualität als solche bewerten, führte aber an, dass viele kleinere Sender bei vielen Sendungen gar keine Untertitelung durchführen würden und dass sie sich bei diesen eine flächendeckendere Untertitelung wünschen würde. Wenn die Frage, welcher Fernsehsender die besten Untertitel zur Verfügung stellt, beantwortet wurde, kam am häufigsten NHK als Antwort, beziehungsweise allgemein die größeren Sender. Dort gäbe es viele untertitelte Sendungen und diesen sei außerdem von der Verzögerung und der Lesbarkeit her einfach zu folgen. Problematisch wären hingegen die Regionalsender. Hier sei vor allem ein Mangel an Untertiteln das Problem. Ähnlich verhalten sich die Antworten bei den Live-Untertiteln, auch hier wird NHK als gutes Beispiel genannt. Besonders lobend erwähnt werden die verschiedenen Farben in der Untertitelung und der Einsatz der Respeaking-Methode.

6.2.3 Formale Kriterien

Einen weiteren Teil der Befragung machten die formalen Kriterien aus. In Bezug auf die allgemeine Lesbarkeit der Untertitel fiel das Urteil überwiegend positiv aus. Alle von mir befragten Personen gaben an, dass die Untertitel gut lesbar seien. T1 meinte insbesondere, dass

die Untertitel *heutzutage* gut lesbar seien, was auf eine Verbesserung der Situation in der letzten Zeit hindeutet. Als Begründung für die gute Lesbarkeit wurde noch angeführt, dass die Untertitel meistens auf dem Bildschirm in ihrer Position angepasst werden, um zu verhindern, dass sie wichtige Teile verdecken. Allerdings merkte T6 an, dass dieses Verdecken dennoch manchmal vorkäme und Probleme machen würde. Besonders hervorgehoben werden die Schriftart und die Positionierung der Untertitel, die passend zur Position des Sprechers erfolgt. Eine Person meinte auch, dass sie diese deswegen einfacher zu lesen fände als jene englischsprachiger Produktionen. Es erfolgt auch ein Vergleich mit dem amerikanischen und europäischen Raum:

欧米の一部の字幕は流れるように動く（スクロール）ことがあるが、日本の字幕は固定表示で視線移動が少なく、読みやすいと思う。

In Amerika und Europa bewegen sich die Untertitel manchmal fließend über den Bildschirm (scrolling), aber in Japan werden sie in einem Stück angezeigt, sodass man die Augen nicht viel bewegen muss. Dadurch sind sie finde ich gut lesbar. (Interview T7:27)

Es herrschte eine leichte Uneinigkeit bezüglich dessen, wie ein Wechsel der SprecherInnen in den Untertitel angezeigt werden soll. Es gab sowohl den Wunsch, den Namen der Person den Untertiteln voranzustellen, also auch jenen, die Personen farblich so zu unterscheiden, dass ein Wechsel der Person daraus unmissverständlich ersichtlich ist. T3 wiederum wünschte sich zum Beispiel sowohl Namen als auch einen Wechsel der Farbe (beide Varianten werden momentan im japanischen Fernsehen bereits so praktiziert). T5 wünscht sich zudem eine Änderung der Schrift mit der Änderung des Sprechers. Es geht nicht klar hervor, ob die befragten Personen zu unterschiedlichen Sendungen unterschiedliche Varianten der Anzeige vorziehen würden. Zusammenfassend war der einzige Kritikpunkt, der genannt wurde, dass die Untertitel manchmal dennoch wichtige Teile des Bildschirms verdecken. Die Sprecher sollen durch Farben gekennzeichnet werden, beziehungsweise auch durch die Verwendung der Namen. Die Live-Untertitelung bei den Lokalsendern sollte außerdem verbessert werden.

Anschließend beschäftigte sich der Fragebogen mit der Geschwindigkeit der Untertitel. In der ersten Gruppe der Interviewten wurde hierbei die Geschwindigkeit der Untertitel als passend empfunden. T1 merkte an, dass die Untertitel in ihrer Geschwindigkeit an jene des Programmes angepasst werden sollten. Auf Fragen der Lesbarkeit in Bezug auf die

Geschwindigkeit nahm sie hierbei scheinbar keine Rücksicht, beziehungsweise erschien ihr diese nicht als Kriterium.

話されていることと同じスピードが良いです。臨場感が大切です。早口なら、字幕の表示も早くあるべきです。

Es sollte dieselbe Geschwindigkeit haben wie der gesprochene Text. Es ist wichtig, dass es realistisch ist. Wenn eine Person schnell spricht, müssen die Untertitel auch schnell angezeigt werden. (Interview T1:58-59)

In der zweiten Gruppe gehen die Meinungen hingegen ein wenig auseinander. Während drei von fünf Personen die Geschwindigkeit für richtig halten, findet sie eine weitere Person zu langsam, noch eine jedoch manchmal zu schnell. Es überwiegt aber der Anteil der Personen, für die die Geschwindigkeit passend ist.

Bei der Thematik der formalen Kriterien zeigt sich wieder einmal die Heterogenität der Zielgruppe, da keine Einigkeit darüber besteht, wie ein Sprecherwechsel am besten zu kennzeichnen ist. Immerhin waren die TeilnehmerInnen relativ durchgehend der Meinung, dass die Geschwindigkeit der Untertitel im Allgemein passend sei.

6.2.4 Wörtlichkeit der Untertitel

Die nächste Frage brachte einige Schwierigkeiten in der Beantwortung mit sich. Gefragt wurde, ob die TeilnehmerInnen durch Editierung gekürzte Untertitel oder wörtliche Untertitel bevorzugen und ob diese Wahl dieselbe bliebe, sollten die wörtlichen Untertitel eventuell mehr Fehler aufweisen als ihre gekürzten Varianten. Die meisten Befragten schienen jedoch trotz der Frage nach den Fehlern, eher die Verzögerung als auftretende Fehler als Problematik im Blick zu haben. Meine Kontaktperson bei der zweiten Organisation bat sogar darum, diese Frage in diese Richtung umformuliert stellen zu dürfen, da sie sie so für realistischer hält. Dies zeigt vermutlich, dass die Geschwindigkeit und Verzögerung eine besonders große Rolle für das Verständnis spielt und wichtiger ist als mögliche Fehler.

Alle drei Personen der ersten Gruppe wünschten sich möglichst wörtliche Untertitel. T1 räumte allerdings ein, dass es auch auf die Art der Sendung ankäme, ob eher die Geschwindigkeit oder die Korrektheit Vorrang haben sollte. Bei der zweiten Gruppe gingen die Meinungen wieder mehr auseinander. Obwohl sich eine leichte Präferenz für vollständige Untertitel erkennen lässt, ist diese nicht besonders ausgeprägt. Wie gesagt muss aber noch erwähnt werden, dass diese Frage bei der zweiten Befragung von der Kontaktperson

umformuliert und eher als eine Frage der Geschwindigkeit aufgefasst wurde. Die TeilnehmerInnen entschieden sich somit für langsame und vollständige, aber dadurch möglicherweise verzögerte Untertitel, anstatt für fehlerbehaftete und unvollständige Untertitel mit wenig Verzögerung. Ob diese gänzlich wörtlich sind, blieb damit nur eine Nebenfrage, während sich der Kern um die Verzögerung drehte. Es scheint also so zu sein, dass eine absichtliche Verkürzung, um den Leseaufwand zu verringern, gar nicht wirklich in Betracht gezogen wird.

Insgesamt kann also gesagt werden, dass die Zielgruppen eher eine wörtliche Untertitelung präferierten, wenngleich sie gewisse Einschränkungen einräumten. Die Wörtlichkeit der Untertitel wird eher als eine Frage der auftretenden Verzögerung als in Zusammenhang mit der Lesegeschwindigkeit gesehen.

6.2.5 Prozess und Zukunft der Live-Untertitelung

Der nächste Teil der Befragung befasste sich mit der Erstellung von Live-Untertiteln und es sollte damit ermittelt werden, inwiefern die Befragten über den Prozess der Live-Untertitelung Bescheid wissen und welche Probleme sie damit assoziieren. Allgemein waren die Kenntnisse zu diesem Prozess sehr gut. Es muss allerdings erwähnt werden, dass die Befragten als Mitglieder einer Gehörlosen-Organisation sich wahrscheinlich stärker mit der Thematik befassen, als man es von der Durchschnittsperson aus der Zielgruppe erwarten kann.

Fast alle TeilnehmerInnen gaben an, den Prozess zur Erstellung der Untertitel zu kennen. Es wurden auch verschiedene Methoden, mit denen Live-Untertitel produziert werden, genannt, wie das Steno-Tastatur und die Spracherkennung, beziehungsweise einige Male auch spezifisch das Respeaking als Methode. Eine Teilnehmerin gab sogar an, bei einer Live-Untertitelung bereits Gast gewesen zu sein. Das Wissen war also sehr detailliert.

Es werden zudem noch folgende weitere Probleme bei der Erstellung von Live-Untertitelung aufgeführt:

- Manchmal werden nicht alle nichtsprachlichen Informationen transportiert.
- Es gibt bei manchen Sendungen keine Untertitel, zum Beispiel auch bei fremdsprachigen Sendungen
- Manchmal ist es schwierig festzustellen, wer in einer Live-Sendung der oder die SprecherIn ist.

Die meisten dieser Probleme lassen sich auf die Verzögerung zurückführen, durch die ein Informationsverlust entsteht.

Den Abschluss der Befragung bildeten zwei Fragen, die etwas allgemeiner formuliert sind. Hierdurch sollten Aspekte angesprochen werden, die die Befragten selbst für wichtig halten, aber die von den vorgegebenen Fragen nicht abgedeckt wurden. Hierdurch sollte verhindert werden, zu sehr an den Meinungen der TeilnehmerInnen vorbei zu fragen.

Die erste Frage befasste sich damit, wie Barrierefreiheit im Fernsehen am besten gewährleistet werden kann. Hierzu gab es unterschiedliche Antworten. T1 sprach die Problematik der Finanzierung an. Zwar würde die Untertitelung bei den großen Fernsehsendern voranschreiten, bei den Regionalsendern sei aber aufgrund mangelnder Finanzierung noch weniger Fortschritt zu bemerken. Dies sollte durch Förderungen der Regierung ausgeglichen werden. Darüber hinaus seien auch mehr Untertitelungsfirmen notwendig. T2 wünscht sich, dass die Untertitel auf Fernsehern an öffentlichen Orten zugeschaltet wären. T3 hingegen hält eine Gebärdensprach-Dolmetschung während der Fernsehsendungen für die beste Möglichkeit, weil dadurch keine Verzögerung entsteht. T1 führt auch noch ein, dass sie sich wünschen würde, die Untertitelung würde auf die Online-Medien ausgedehnt, anstatt nur im Fernsehen stattzufinden.

Eine häufig auftretende Meinung zur Frage nach der Barrierefreiheit war auch, dass es wichtig sei, darüber nachzudenken, wie man möglichst viele Menschen zugleich in eine Sendung einbinden und diese diesen zugänglich machen könnte. Hierzu wurde die Wichtigkeit von Wahlmöglichkeiten unterstrichen, und universelles Design als Ausgangspunkt anstatt der Barrierefreiheit genannt.

Netflixのように、すべてのコンテンツに対して字幕付与がスタンダードであり、言語も自由に選択できること。バリアフリーの観点ではなく、ユニバーサルデザインの概念で字幕付与普及を進めてほしい

Ich wünsche mir, dass wie bei Netflix alle Inhalte standardmäßig mit Untertiteln versehen sind und dass man die Sprache frei wählen kann. Ich möchte außerdem, dass es eher von dem Standpunkt eines universellen Designs betrachtet wird als von jenem der Barrierefreiheit. So lässt sich die Untertitelung fördern. (Interview T6:33)

Ebenfalls erwähnt wurde erneut die Möglichkeit der absichtlichen Verzögerung zwischen Aufzeichnung und Ausstrahlung.

Anschließend wurden die Personen noch einmal gefragt, ob sie etwas hinzuzufügen hätten. Von den meisten kam dazu keine weitere Angabe. Eine Person hingegen schrieb ihre

Meinung sehr ausführlich nieder. Ihr Anliegen war hierbei, dass viele Leute sich der Wichtigkeit von Untertitelung nicht bewusst seien. Obwohl Japan eine alternde Gesellschaft sei und viele Menschen Untertitel brauchen würden, wüssten sie häufig nicht, wie diese funktionieren. Zudem sei die mangelnde Untertitelung in parlamentarischen Debatten ein besonders großes Problem, da es die politische Partizipation behindere. Es brauche mehr Aufmerksamkeit in der Gesellschaft für diese Probleme. Auch eine andere Person beklagt die zu geringe Teilnahme:

テレビ CM 部分はシステム上、字幕をつけにくかった。改善の動きがあるがまだまだである。広告会社は聴覚障害の消費者をみすみす見逃している。これまでに当事者不在でシステムを作ってきた弊害がここにも出ている。

Es ist aufgrund des Systems schwierig, Untertitel zu Werbungen hinzuzufügen. Es gibt schon Ansätze, dies zu ändern, aber es braucht noch viel mehr. Die Unternehmen schließen damit gehörlose Menschen als potentielle Kundschaft aus. Dies zeigt auch die Problematik eines Systems, welches entwickelt wurde, ohne die Zielgruppe miteinzubeziehen. (Interview T7:34)

Die für diese Arbeit befragten Personen wussten im Allgemeinen sehr gut über den Ablauf der Untertitelung Bescheid und konnten auch verschiedene Methoden nennen und teilweise beschreiben. Die größten Probleme bei der Live-Untertitelung treten durch die Verzögerung und durch falsch umgewandelte Schriftzeichen auf. Generell wünschen sich die Teilnehmerinnen mehr Sichtbarkeit und Bewusstsein für die Untertitelung sowie mehr staatliche Unterstützung, um diese flächendeckend zu gewährleisten.

7 Schlussfolgerungen und Ausblick

Diese Arbeit befasste sich mit verschiedenen Aspekten der Live-Untertitelung in Japan. Ein besonderer Fokus lag hierbei auf der Respeaking-Methode. Obgleich die Forschung zum Thema Live-Untertitelung in den letzten Jahren stark zugenommen hat, ist es immer noch schwierig, Arbeiten zu deren Anwendung in Japan zu finden. Insbesondere sind die meisten Informationen nur in japanischer Sprache zugänglich. Diese Arbeit sollte daher einen Beitrag dazu leisten, diese Informationslücke zu schließen. Zu diesem Zweck wurde sowohl eine Analyse der vorhandenen japanischen Arbeiten als auch eine Befragung relevanter Personen bei NHK sowie aus den Zielgruppen vorgenommen. Durch diese Erhebung konnte ein grundsätzliches Bild der Untertitelung in Japan, insbesondere bei NHK, ermittelt werden. Im Folgenden werden die wichtigsten Erkenntnisse erneut zusammengefasst.

Grundsätzlich wird in Japan seit Anfang der 2000er-Jahre mit der Live-Untertitelung im Fernsehen experimentiert. Anfangs wurde mit direkter Spracherkennung gearbeitet, später dann mit einer Steno-Tastatur, ehe sich schließlich die Respeaking-Methode etablierte. Heute sind etwa 98 Prozent des gesamten japanischen Fernsehprogramms untertitelt. Die Untertitelungsarbeit findet sowohl bei NHK als auch bei anderen Untertitelungsfirmen mit deren eigenen Untertitelungsprogrammen statt. In beiden beforschten Abteilungen wird außerdem im Team gearbeitet, wobei ein/e MitarbeiterIn für das Respeaking zuständig ist und eine variable Anzahl weiterer MitarbeiterInnen die Korrektur der Untertitel übernehmen. Aufgrund der hohen Komplexität der japanischen Sprache wird es nicht als möglich erachtet, beide Rollen in einer Person zu vereinen.

Als besonders wichtige Fähigkeiten für das Respeaking werden eine klare Aussprache, sowie die Fähigkeit, Sätze im Kopf (um)formulieren zu können, erachtet. Viele MitarbeiterInnen waren vor ihrem Eintritt in die Live-Untertitelung als NachrichtensprecherInnen tätig, was in Bezug auf diese Fähigkeiten von Vorteil ist. Für die Korrektur sind besonders gute Schriftzeichenkenntnisse und generelle Sprachkenntnisse wichtig. Da bei NHK die beiden Aufgaben getrennt sind, müssen die MitarbeiterInnen nicht sowohl die für RespeakerInnen als auch die für EditorInnen notwendigen Kompetenzen mitbringen. Es wird keine spezielle Ausbildung vorausgesetzt und die Eignung wird in einem normalen Bewerbungsgespräch überprüft. Sowohl bei NHK als auch bei Faith findet die Ausbildung *in-house* statt und ist sehr praktisch orientiert.

Es gibt keine staatlich vorgegebenen Richtlinien bezüglich qualitativer Standards, sondern diese beschreiben nur einen quantitativen Anspruch an die Anzahl der Sendungen, die

untertitelt werden sollten. Sowohl bei NHK als auch bei Faith werden die Untertitel regelmäßig durch eine dazu beauftragte Person hinsichtlich ihrer Genauigkeit und Qualität überprüft. Außerdem müssen die RespeakerInnen eine gewisse Genauigkeit in Prozent erreichen, ehe sie ihre erste Live-Sendung untertiteln dürfen. Genauigkeit wird hier in Bezug auf das Respeaking allerdings rein phonetisch gemessen, da es die Aufgabe der EditorInnen ist, Homonyme den korrekten Schriftzeichen zu zuordnen. Die Verzögerung kann je nach Sendungsart stark variieren, liegt aber durchschnittlich etwa bei sechs bis zehn Sekunden. Mit einer absichtlichen Verzögerung zwischen Aufzeichnung und Ausstrahlung wird im Allgemeinen nicht gearbeitet, auch wenn es schon einmal Experimente hierzu gab.

Es wird durchaus Wert auf eine hohe Wörtlichkeit der Untertitel gelegt und nur editiert, wenn dadurch kein Informationsverlust entsteht oder eine vollständige Wiedergabe aufgrund von Gefahr einer Verzögerung nicht möglich ist. Die Lesegeschwindigkeit der Zielgruppen wird grundsätzlich als hoch genug eingeschätzt, um auch schnelleren Untertiteln folgen zu können.

Feedback von den Zielgruppen erhält NHK bei Konferenzen mit Vertretern der Gehörlosen-Vereinigungen. Hier sehen sie sich mit den unterschiedlichen Erwartungen der Zielgruppen konfrontiert. Allgemein wird das Feedback in Bezug auf die Untertitelung aber als eher positiv eingestuft. Die Zukunft der Live-Untertitelung sehen alle Befragten in der Spracherkennung, auch wenn diese momentan noch mit Problemen behaftet ist. Tatsächlich wurde die Meinung geäußert, dass der Beruf des Respeakers eventuell nicht mehr notwendig sein wird, auch wenn der Output von KI-gestützten Programmen wahrscheinlich noch von Menschen kontrolliert wird werden müssen.

Bei der Zielgruppenbefragung fanden sich, wie zu erwarten gewesen war, verschiedene Meinungen dazu, wie Untertitelung aussehen sollte. Generell ist eine Ausdehnung der Untertitelung natürlich gewünscht, die Beschränkung auf 07:00 bis 24:00 Uhr wird als zu eingeengt empfunden. Obwohl die Untertitelungsrate bei großen, japanweit ausgestrahlten Sendern und insbesondere bei NHK relativ hoch ist, fehlen die Untertitel bei Regionalsendern häufig noch.

Die Qualität der Untertitelung wird auch im Live-Bereich als eher hoch eingestuft. Das am häufigsten auftretende Problem ist hierbei die Verzögerung. Außerdem wünschten sich die TeilnehmerInnen eine bessere Kennzeichnung der Untertitel, eine vermehrte Miteinbeziehung der Zielgruppen in den Prozess und mehr Unterstützung für die UntertitlerInnen. Die Lesbarkeit der Untertitel und die Geschwindigkeit wird allgemein als passend empfunden, auch wenn es hierzu auch einige Anmerkungen und Verbesserungsvorschläge gab. Es gibt keine Einigkeit

darüber, wie ein Sprecherwechsel am besten darzustellen ist, sowohl Farben als auch Namen werden von unterschiedlichen Personen präferiert. Wörtliche Untertitel werden generell den editierten vorgezogen.

Für die Zukunft der Untertitelung wünschen sich die TeilnehmerInnen meiner Umfrage vor allem mehr Anerkennung für die Untertitelung. Auf Bildschirmen an öffentlichen Plätzen sollten Untertitel eingeblendet werden, und generell sollte das Bewusstsein über die Wichtigkeit der Untertitelung in der Bevölkerung gestärkt werden. Es wird auch mehr finanzielle Unterstützung von Seiten der Regierung erwünscht, um die Barrierefreiheit im Fernsehen zu fördern. Dazu sollten auch Gebärdensprach-Sendungen weiter ausgebaut werden.

Generell ist Japan seit langer Zeit ein Land, das häufig Untertitel einsetzt, was vielleicht zur frühen Adaption und Akzeptanz in der Untertitelung beiträgt. Selbst bei Fernsehsendungen ohne besondere barrierefreie Anpassungen werden häufig *open captions* verwendet. Obwohl die Live-Untertitelung bereits weit fortgeschritten ist und mit verschiedenen Methoden gearbeitet wird, gibt es dennoch weiterhin Schwierigkeiten, die überwunden werden müssen. In mancher Hinsicht sind die Problematiken der Live-Untertitelung natürlich dieselben wie auch in anderen Sprachgebieten. Die Live-Situation mit der entstehenden Verzögerung und die unterschiedlichen Erwartungshaltungen der Zielgruppen sind eine Herausforderung, die überall zu finden ist. Zugleich bringt die japanische Sprache durch das komplizierte Schriftsystem und die häufigen Homonyme aber auch einzigartige Anforderungen mit sich. Diese werden momentan vor allem durch einen erhöhten Personalaufwand bei der Korrektur bewältigt.

7.1 Mögliche weitere Forschungsansätze

Diese Arbeit soll dazu dienen, einen ersten Überblick zur Thematik Live-Untertitelung und der Verwendung der Respeaking-Methode in Japan zu bieten. Da hierzu im deutschsprachigen Raum bisher noch kaum geforscht wurde, bieten sich noch viele Möglichkeiten, weiter in die Tiefe gehende Forschungen durchzuführen. Tatsächlich wäre es, um ein ganzheitliches Bild zeichnen zu können, empfehlenswert, eine Beobachtung der Live-Untertitelung bei NHK durchzuführen, um die alltäglichen Arbeitsabläufe kennen zu lernen. Es hätte auch ein solches Angebot für diese Forschung bestanden, das aber aufgrund der Reiserestriktionen 2020 und 2021 nicht wahrgenommen werden konnte.

Es gäbe auch viele Themen, mit denen sich Forschende noch im Detail beschäftigen könnten. Zum Beispiel könnte näher auf die Problematik der japanischen Sprache in

Zusammenhang mit Methoden der Spracherkennung eingegangen werden. Beispielsweise könnten hier Untertitelungstexte vor und nach der Korrektur durch EditorInnen analysiert oder die einzelnen Schwierigkeiten näher beschrieben werden.

Ein weiterer interessanter Punkt wäre die Kooperation im Team zwischen den RespeakerInnen und EditorInnen, die ja anders abläuft als in vielen anderen Ländern. Auch Themen wie die Ausbildung und Hintergründe der handelnden Personen, das Qualitätsmanagement und die Interaktion mit den Zielgruppen erlauben Fragestellungen, die in dieser Arbeit unbeantwortet geblieben sind. Ebenso wäre ein Vergleich mit der Untertitelung in einem europäischen Land möglich.

In Zukunft wird Japan mit Sicherheit ein interessantes Forschungsgebiet im Bereich der Untertitelung sein, da durch die Überalterung der Bevölkerung ein großer Bedarf an Untertitelung für schwerhörige und gehörlose Personen besteht. Darüber hinaus stellt die japanische Sprache eine spezielle Herausforderung dar, mit der durch eigene Methoden umgegangen werden muss. Daher wird es interessant sein, zu beobachten, wie japanische Fernsehanstalten in Zukunft mit diesen Problematiken umgehen werden, um eine umfangreichere Live-Untertitelung zu gewährleisten.

Literatur

- Arumí Ribas, Marta & Romero Fresco, Pablo (2008). A Practical Proposal for the Training of Respeakers. *Journal for Specialised Translation* 10, 106–127.
- BMASGK (2016). *UN-Behindertenrechtskonvention Übereinkommen über die Rechte von Menschen mit Behinderungen und Fakultativprotokoll. Neue deutsche Übersetzung*. Wien: Bundesministerium für Arbeit, Soziales, Gesundheit und Konsumentenschutz <https://broschuerenservice.sozialministerium.at/Home/Download?publicationId=19> (Stand: 19.06.2021)
- Bogner, Alexander & Littig, Beate & Menz, Wolfgang (2014). *Interviews mit Experten. Eine praxisorientierte Einführung*. Wiesbaden: Springer.
- Chaume, Frederic (2013). The turn of audiovisual translation: New audiences and new technologies. *Translation Spaces* 2, 105–123.
- Chmiel, Agnieszka & Lijewska, Agnieszka & Szarkowska, Agnieszka & Dutka, Łukasz (2017). Paraphrasing in respeaking – comparing linguistic competence of interpreters, translators and bilinguals. *Studies in Translation Theory and Practice* 26 (5), 725–744.
- Geens, Rita (2008) Live Subtitling through Respeaking: a new Discipline in Interpretation? Translation Conference 2008. *The Changing Face of Translation*, 100–110.
- De Castro, Mercedes & Puente Rodríguez, Luis & Ruiz Mezcua, Belén (2015). Synchronized Subtitles in Live Television Programmes. In: Piñero, Rocío Baños & Díaz Cintas, Jorge (Hg): *Audiovisual Translation in a Global Context. Palgrave Studies in Translating and Interpreting*. London: Palgrave Macmillan, 51–71.
- Díaz Cintas, Jorge (2009). *New Trends in Audiovisual Translation*. Bristol: Blue Ridge Summit. Multilingual Matters.
- Díaz Cintas & Baños Piñero, Rocío (2016). *Audiovisual Translation in a Global Context: Mapping an Ever-changing Landscape*. London: Palgrave Macmillan.
- Díaz Cintas, Jorge & Remael, Aline (2014). *Audiovisual Translation: Subtitling*. London/New York: Routledge.
- Díaz Cintas, Jorge & Neves, Joslia & Matamala, Anna (2010). *New Insights into Audiovisual Translation and Media Accessibility: Media for All 2*. New York: Rodopi.
- Díaz Cintas, Jorge & Orero, Pilar & Remael, Aline (2007). *Media for all: a global challenge*. Amsterdam/New York: Rodopi.
- Eugeni, Carlo (2008). Respeaking the TV for the Deaf: For a Real Special Needs-Oriented Subtitling. *Studies in English Language and Literature* 21, 37–47.
- European Commission (2021) *European accessibility act*. <https://ec.europa.eu/social/main.jsp?catId=1202&langId=de> (Stand: 19.06.21).
- Fresno, Nazaret (2018). Of bad hombres and nasty women; the quality of the live closed captioning in the 2016 US final presidential debate. *Perspectives Studies in Translatology* 27(1), 1–17.
- Guillot, Marie-Noëlle (2018). Subtitling on the cusp of its futures. In: Perez-Gonzales, Luis (Hg.) *The Routledge Handbook of Audiovisual Translation*. Abingdon: Routledge, 31–47.
- 本間真一 Homma, Shinichi (2010) 字幕製作のための音声認識 [Spracherkennung zur Untertitelung] NHK 技研 RandD 122, 25–31.
- 本間真一 Homma, Shinichi & 尾上和穂 Onoe, Kazuo & 小林彰夫 Kobayashi, Akio & 佐藤庄衛 Satou, Shoei & 今井享 Imai, Toru & 都木徹 Takagi, Tohru (2007). 番組音声とリスピーク音声の自動認識を利用したニュース字幕制作実験. [Experiment of

- Real-Time Captioning for Broadcast News Using Speech Recognition of Direct Program Sound and Respoken Utterance] 年映像情報メディア学会年次大会.
- 本間真一. Homma Shinichi & 小林彰夫 Kobayashi, Akio & 奥貴裕 Oku, Takahiro & 佐藤庄衛 Satou, Shoei & 今井享 Imai, Toru & 都木徹 Takagi, Tohru (2009). ダイレクト方式とリスピーク方式の音声認識を併用したリアルタイム字幕制作システム [Real-Time Closed Captioning System Using Speech Recognition of Direct Program Sound and Re-Spoken Utterances]. 映像情報メディア学会誌 63(3), 331–338.
- Hopf, Christel (2019.) Qualitative Interviews in der Sozialforschung. Ein Überblick. In: Flick, Uwe & Kardorff, Ernst von & Steinke, Ines (Hgs) *Qualitative Forschung : ein Handbuch*. 13. Aufl. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag, 177–182.
- 伊藤崇之 Ito, Takayuki (2013) メディアアクセシビリティの現状と展望 [Momentane Situation und Zukunft der Barrierefreiheit in den Medien]. 映像情報メディア学会誌 67(11), 949 – 954.
- ITU (2018). *Multimedia services Accessibility terms and definitions*. Telecommunication standardization sector of ITU <https://www.itu.int/itu-t/recommendations/rec.aspx?rec=12624>
- Japanese Federation of the Deaf (2021). 連盟について [Über die Organisation] <https://www.jfd.or.jp/about> (Stand 10.04.21)
- Karl-Franzens-Universität Graz (2017). Curriculum für das Masterstudium Übersetzen. https://online.uni-graz.at/kfu_online/wbMitteilungsblaetter_neu.display?pNr=14881&pDocNr=3772040&pOrgNr=14190 (Stand: 19. 06. 2021)
- Lamnek, Siegfried (2010). *Qualitative Sozialforschung*. 5. Aufl. Weinheim: Beltz Verlag.
- Luyckx, Bieke & Delbeke, Tijs & Van Waes, Luuk & Leijten, Mariëlle & Remael, Aline (2013). Live Subtitling with Speech Recognition. Causes and Consequences of Text Reduction. *Across Languages and Cultures* 14(1), 15–46.
- Mayring, Philipp (2010). *Qualitative Inhaltsanalyse*. Weinheim: Beltz Verlagsgruppe.
- Ministerium für Gesundheit, Wohlfahrt und Arbeit (2018). 平成 28 年生活のしづらさなどに
関する調査 [Umfrage zu Einschränkungen im täglichen Leben aus dem Jahr 2016]. https://www.mhlw.go.jp/toukei/list/dl/seikatsu_chousa_b_h28.pdf [Stand 19.06.21]
- Munday, Jeremy (2016). *Introducing Translation Studies. Theories and applications*. 4. Aufl. New York: Routledge.
- Orero, Pilar & Matamala, Anna (2010). *Listening to Subtitles. Subtitles for the Deaf and Hard of Hearing*. Bern: Peter Lang.
- O’Sullivan, Carol & Cornu, Jean-François (2018). History of Audiovisual Translation. In Perez-Gonzales, Luis (HG.) *The Routledge Handbook of Audiovisual Translation*. Abingdon: Routledge, 31–47.
- Pérez-González, Luis (2018). *The Routledge Handbook of Audiovisual Translation*. Oxon & New York: Taylor and Francis.
- Remael, Aline (2007). Sampling subtitling for the deaf and the hard-of-hearing in Europe. In: Díaz Cintas, Jorge; Orero, Pilar & Aline Remael (Hg.) *Media for All. Subtitling for the Deaf, Audio Description, and Sign Language*. Amsterdam und New York: Rodopi, 23–52.
- Remael, Aline & Reviers, Nina & Vandekerckhove, Reinhild. (2016). From Translation Studies and audiovisual translation to media accessibility: Some research trends. *Target* 28(2), 248–260.

- Robert, Isabelle S. & Schrijver, Iris & Diels, Ella (2019). Live subtitlers: Who are they? A survey stud. *Linguistica Antverpiensia. New Series: Themes in Translation Studies* 18, 101–129.
- Romero-Fresco, Pablo (2009). More haste less speed: Edited versus verbatim respoken subtitles. *Vigo International Journal of Applied Linguistics* 6 (1), 110–133.
- Romero-Fresco, Pablo (2011). *Subtitling through Speech Recognition: Respeaking*. Manchester : St. Jerome Publ.
- Romero-Fresco, Pablo (2012) Respeaking in Translator Training Curricula. *The Interpreter and Translator Trainer* 6(1), 91–112.
- Romero-Fresco, Pablo (2015). *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe*. Bern: Peter Lang AG.
- Romero-Fresco, Pablo (2016) Accessing communication: The quality of live subtitles in the UK. *Language & Communication* 49, 56–69.
- Romero-Fresco, Pablo (2018) Reception studies in live and pre-recorded subtitles for the deaf and hard of hearing. In: Di Giovanni, Elena & Gambier, Ives (Hgs.) *Reception Studies and Audiovisual Translation*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 199–224.
- Romero-Fresco, Pablo & Martínez Perez, Juan (2015). Accuracy Rate in Live Subtitling: The NER Model. In: Baños Piñero, Rocío; Díaz Cintas, Jorge (Hg.) *Audiovisual Translation in a Global Context. Mapping an Ever-changing Landscape*. London: Palgrave Macmillan, 28–50.
- Romero-Fresco, Pablo & Pöchhacker, Franz (2017). Quality assessment in interlingual live subtitling: The NTR model. *Linguistica Antverpiensia New Series: Themes in Translation Studies* 16, 149–167.
- 佐藤 庄衛 Satou, Shouei (2018). 聴覚障害者のための生放送番組への字幕付与システム 佐藤 [Untertitelungssystem für Live-Sendungen für Personen mit Hörbehinderungen] 日本音響学会誌 74 (3), 163–170.
- Universitat Autònoma de Barcelona (2020). 2020/2021 Superimposed Voices and Respeaking B-A https://guies.uab.cat/guies_docents/public/portal/html/2020/assignatura/44385/en (Stand: 01. 06. 2021)
- University of Roehampton (2020). *Audiovisual Translation – Postgraduate*. <https://www.roehampton.ac.uk/postgraduate-courses/audiovisual-translation/> (Stand: 28. 1. 2020)
- Universität Wien (2018). *ufind: UE Respeaking*. <https://ufind.univie.ac.at/en/course.html?lv=340418&semester=2018W> (Stand: 28. 1. 2020)
- Walter, Cornelia (2012). *Respeaking - Intralinguales Simultandolmetschen für die Untertitelung*. Masterarbeit, Universität Wien.
- Zennancho (2021). 概要 [Übersicht] <https://www.zennancho.or.jp/info/summary/> (Stand: 10.04.21)

Anhang

Anhang 1 Fragebögen

Fragebogen Untertitelung Japanisch

総合的な質問

1. NHK では生放送に字幕を付与することをいつから開始しましたか。生放送字幕の歴史を簡単に説明してください。
2. 現在では生放送の何パーセントが字幕放送ですか。
3. 生字幕のためのリスピーク方式の使用はいつから開始しましたか
4. リスピーク以外どんな方式が使用されていますか。
5. リスピーク方式はどんな種類の番組で使用されていますか。
6. 音声認識のためにどんなプログラムが使用されていますか。
7. そのプログラムは誰によって開発されましたか。そのプログラムを使用する決定はどのように行われましたか。
8. 生字幕は日本語から日本語にだけですか。ほかの言語から日本語に翻訳していた生字幕はありますか。また、将来的にそういった字幕を作る計画はありますか。
9. リスピークは何人によって行われますか。そして、仕事はどんな役割に分けていますか。
10. 校正者の全員がリスピークの仕事もしますか。そして逆にリスピークをする人が校正の仕事をすることはありますか。
11. 放送中役割を交代しますか。休憩はありますか。交代はどのくらいの間隔で行われますか。
12. 新しい人員を探す時重視される能力や資格はなんですか。
13. 字幕の品質はどのようにチェックしますか。そのためのガイドラインなどはありませんか。
14. 字幕の精度についてどういう基準がありますか。
15. 形式について基準はありますか。話者の交代はどのように示されていますか。
16. 字幕の遅延について規定はありますか。どのぐらいの字幕の遅延が許容範囲内とされていますか？ 字幕が同時に表示されるように、番組の放送を意図的に遅らせている国がありますが、これは日本でも考慮されましたか。
17. 句読点は字幕で使われていますか。それは、だれが挿入しますか。
18. 字幕の速さと正しさのどちらがより大切だと思いますか。
19. 元の言葉通りの字幕を作ることはどの程度重要視されていますか。間違いを避けるためにテキストを省略することはよくありますか。これについて規定はありますか。

20. 字幕の対象者、例えば聴覚障害者や高齢者とのコミュニケーションが行われたり、そういった人からのフィードバックを貰ったりすることはありますか。
21. 生字幕をさらに充実させる予定はありますか。生字幕の未来についてどう思いますか。

リスピーカーへの質問

1. いつリスピーカーとして働き始めましたか。
2. 字幕キャスターになる以前の職歴・学歴を教えてください。
3. リスピーク方式の字幕キャスターのための教育はどこで受けましたか。
4. その教育はどういった内容でしたか。そしてどのくらい時間かかりましたか。
5. その教育についてどう思いますか。リスピークをするためにいい準備だと思いますか。もっと詳しい説明や教育があったらいいと思う部分がありますか。
6. どんな種類の番組の字幕付与をしていますか。
7. リスピーク方式はどのくらいの頻度で使いますか。例えば一か月や一週間に何回。
8. 生字幕以外にほかの仕事もありますか。
9. 就業時間はどのくらいですか。
10. あなたの考える一番よくある問題と間違えの原因はなんですか。
11. 技術的な問題も発生しますか。どのくらいの頻度でありますか。
12. こういう問題や間違いを避けるためにどの戦略を使用しますか。
13. どのようにストレスに対処しますか。
14. あなたの考える字幕を付与する際の一番難しい状況はどのような状況ですか。
15. 同音異義語や同じ読み方がある漢字について、何割くらいはシステムが自動的に修正できますか。そして何割くらいは校正者の責任になりますか。
16. 字幕の何割くらいが元の言葉通りだと思いますか。
17. リスピーカーにとって一番大切な資格はどれだと思いますか。
18. リスピーカーと通訳者にとっての必要な能力は似ていると思いますか。
19. 放送前にリスピークのための準備はありますか。どういう準備ですか。
20. 日本におけるリスピークの未来についてどう思いますか。

Fragebogen Untertitelung Deutsch

Allgemeine Fragen:

1. Seit wann wird bei NHK Untertitelung für Live-Sendungen eingesetzt? Erzählen Sie vielleicht kurz die Geschichte:
2. Wie viel Prozent der Live-Sendungen werden heute Untertitelt?
3. Seit wann wird Respeaking zur Live-Untertitelung eingesetzt?
4. Welche anderen Methoden werden auch noch eingesetzt?
5. Für welche Arten von Sendungen wird Respeaking eingesetzt?
6. Welches Programm wird zur Spracherkennung eingesetzt?

7. Wer hat diese Programme entwickelt? Wie ist es zur Entscheidung gekommen dieses zu verwenden?
8. Wird nur intralingual untertitelt oder auch interlingual? Wenn nicht, gibt es Pläne für die Zukunft?
9. Wie ist die Rollenverteilung beim Respeaking, wie viele Personen sind beteiligt?
10. Arbeiten alle Editoren auch als Respeaker und umgekehrt?
11. Werden die Rollen gewechselt bzw Pausen gemacht? Ungefähr in welchem Intervallen?
12. Welche Kriterien sind wichtig, wenn neue MitarbeiterInnen eingestellt werden?
13. Wie überprüfen Sie generell die Qualität der produzierten Untertitel und gibt es Richtlinien dazu?
14. Welche Standards für Genauigkeit der Untertitel kommen zum Einsatz?
15. Welche Anforderungen gibt es an das Format? Wie werden Sprecherwechsel angezeigt?
16. Gibt es interne Vorschriften über den Delay? Wenn ja, welcher Delay gilt als akzeptabel? In manchen Ländern (z.B. Belgien) wird absichtlich die Ausstrahlung verzögert damit die Untertitel zur selben Zeit kommen – wurde das bei Ihnen angedacht?
17. Werden in den Untertiteln Satzzeichen verwendet, und wenn ja, wer fügt diese ein?
18. Würden Sie sagen, dass im Zweifelsfall eher die Geschwindigkeit wichtiger ist oder die Korrektheit der Untertitel?
19. Wie viel Wert wird auf Wörtlichkeit gelegt? Wird häufig gekürzt, um die Korrektheit zu erhalten? Gibt es dazu eine genaue Vorschrift?
20. Wie findet Kommunikation mit den Zielgruppen (zb Gehörlose, ältere Menschen) statt?
21. Ist ein Ausbau des Live-Untertitelungsangebots geplant? Wie sehen sie die Entwicklung in der Zukunft?

Fragen an RespeakerInnen:

1. Seit wann arbeiten Sie mit Respeaking?
2. In welchem Bereich haben Sie vorher gearbeitet bzw. welche Ausbildung haben Sie?
3. Wo haben Sie Live-Untertitelung (mit Respeaking) gelernt?
4. Was war der Inhalt der Ausbildung? Wie lange dauerte diese Ausbildung?
5. Wie hilfreich fanden Sie die Ausbildung? Gäbe es Bereich auf die man noch genauer hätte eingehen sollen?
6. Welche Sendeformate untertiteln Sie? Welche Themengebiete?
7. Wie oft in etwa (zb in einer Woche? Monat?) kommt es vor, dass sie Respeaking zur Untertitelung einsetzen?
8. Haben Sie außerhalb der Live-Untertitelung noch andere Aufgaben?
9. Wie sehen Ihre Arbeitszeiten aus?
10. Wo liegen Ihrer Meinung nach die Hauptprobleme /Fehlerquellen beim Respeaking?
11. Sind technische Probleme auch ein Faktor und wenn ja, wie häufig?
12. Welche Strategien verwenden Sie um diese Probleme/Fehler zu vermeiden?
13. Wie gehen Sie mit Stress beim Untertiteln um?
14. Was sind für Sie die schwierigsten Situationen beim Respeaking?
15. Wie sieht es aus mit Homonymen/Kanji mit gleichen Lesungen? Wie viel davon ungefähr löst das System und wie viel müssen Sie auswählen?
16. Wie wörtlich sind die Untertitel Ihrer Meinung nach?

17. Welche Qualifikationen muss ein Untertitler Ihrer Meinung nach vor allem mitbringen?
18. Denken Sie es gibt einen Zusammenhang zwischen den Fähigkeiten die Respeaker haben müssen und jenen von Dolmetschern?
19. Gibt es eine Vorbereitung vor der Sendung für Respeaking? Woraus besteht sie?
20. Wie sehen Sie die Zukunft des Respeakings in Japan?

Fragebogen Zielgruppe Japanisch

1. お名前
2. 性別
3. 年齢
4. 職業
5. 難聴のレベルは何ですか。
6. 何歳のころに聴力を失いましたか。
7. 人工内耳か補聴器を使用していますか。使用すると聴力はどうなりますか。
8. あなたの母語は何語ですか。例えば、日本手話か、日本語か、もしくはどちらも母語だと認識していますか。
9. テレビを見る時に字幕を使用しますか。
10. どのくらいの頻度で字幕を使用していますか。例えば、一か月〇〇回、一週間〇〇回、毎日...
11. どの種類の番組で字幕を使用しますか。字幕が必要ではない種類の番組もありますか。
12. 生放送でも字幕を使用しますか。(例えばニュースやスポーツやほかの事前に字幕を作ることが出来ない放送)
13. あなたの考える、テレビを聴覚障害者にとってバリアフリーにするための一番良い方法は何ですか。
14. 日本のテレビの字幕の量についてどう思いますか。足りですか。足りませんか。
15. 日本のテレビの字幕全般の質についてどう思いますか。どのように改善できますか。
16. リアルタイム字幕全般の質についてどう思いますか。どのように改善できますか。
17. 字幕を見る時、何か問題はありますか。どのような問題ですか。
18. 字幕について何か変えてほしいことはありますか。どういうことですか。
19. 良い字幕を作るために大切なことはどんなことですか。
20. 日本のテレビ局を比較して、ほかより良い字幕を製作しているテレビ局はあると思いますか。それはなぜですか。
21. 日本のテレビ局を比較して、ほかより良いリアルタイム字幕を製作しているテレビ局があると思いますか。どうしてですか。

22. 日本のテレビで出る字幕は読みやすいと思いますか。読みにくいと思いますか。それはなぜですか。
23. 字幕で話者の交代を表す時どうすれば良いと思いますか。
24. 字幕の速さについてどう思いますか。速すぎますか。遅すぎますか。ちょうど良いですか。
25. 話されたことが全部書いている字幕と、修正したり略したりしている字幕のどちらが良いとおもいますか。話されたことが全部書いてある字幕が良いと思う場合、もしその方が修正の入っている字幕よりも間違いが多いとしても、そちらが良いと思いますか。
26. リアルタイム字幕がどのように製作されているということをご存じですか。どのように製作されていると思えますか。
27. リアルタイム字幕を製作する時にどのような問題があると思いますか。
28. ほかに付け加えたいことはありますか。

Fragebogen Zielgruppe Deutsch

1. Name
2. Geschlecht
3. Alter
4. Beruf
5. Wie stark ist Ihr Hörverlust?
6. In welchem Alter haben Sie Ihr Gehör verloren?
7. Nutzen Sie Hörgeräte oder ein Cochlea-Implantat? Wenn ja, wie verändert dies Ihre Hörfähigkeit?
8. Welche würden Sie als Ihre Muttersprache bezeichnen? Japanisch, Japanische Gebärdensprache oder beides zu gleichen Anteilen?
9. Verwenden Sie Untertitel, wenn Sie fernsehen?
10. Wie häufig verwenden Sie Untertitel. Zum Beispiel wie oft in der Woche, wie oft im Monat oder täglich...
11. Für welche Art von Sendungen verwenden Sie Untertitel? Gibt es auch Sendungen, für die Sie keine Untertitel brauchen?
12. Verwenden Sie auch für Live-Sendungen Untertitel?
13. Was denken Sie ist die beste Methode, um das Fernsehen barrierefrei für Gehörlose zu machen?
14. Was denken Sie über die Menge an Untertitelten Sendungen im japanischen Fernsehen? Ist diese ausreichend oder nicht?
15. Was denken Sie generell über die Qualität der Untertitel im japanischen Fernsehen? Wie könnte man diese verbessern?
16. Was denken Sie generell über die Qualität der Live-Untertitel im japanischen Fernsehen? Wie könnte man diese verbessern?
17. Gibt es Probleme bei der Nutzung von Untertiteln? Wenn ja, welche Probleme?
18. Welche Veränderungen würden Sie sich bei Untertiteln wünschen?
19. Was ist wichtig für gute Untertitel?

20. Wenn Sie japanische Fernsehsender vergleichen, gibt es welche, die bessere Untertitel machen? Warum sind diese besser?
21. Wenn Sie japanische Fernsehsender vergleichen, gibt es welche, die bessere Live-Untertitel machen? Denken Sie, dass die Untertitel im Fernsehen gut oder schlecht zu lesen sind? Wieso?
22. Sind Untertitel im japanischen Fernsehen eher einfach oder schwierig zu lesen? Wieso?
23. Wie sollte ein Sprecherwechsel in den Untertiteln angezeigt werden?
24. Was denken Sie über die Geschwindigkeit von Untertiteln? Sind diese eher zu schnell oder zu langsam, oder genau richtig?
25. Bevorzugen Sie Untertitel, die den gesamten gesprochenen Text wiedergeben, oder editierte Untertitel? Wenn Sie vollständige Untertitel bevorzugen, würden Sie dies auch dann noch tun, falls diese mehr Fehler beinhalten?
26. Wissen Sie, wie Live-Untertitel hergestellt werden? Beziehungsweise, was denken Sie, wie diese hergestellt werden?
27. Welche Probleme denken Sie können bei der Erstellung von Live-Untertiteln auftreten?
28. Gibt es noch etwas, dass sie hinzufügen möchten?

Anhang 2 Interviewtranskripte

Interview NHK

- 1 NHK1: 一番目のことですが、生放送に字幕をスタートしたのが 2000 年です。 #00:05:41-6#
- 2 Interviewer: はい, 2000 年。 #00:05:46-5#
- 3 NHK1: 2000 年にですね、ニュース 7 という、7 は七ということなんですけども、そのニュー
- 4 ス番組で音声認識を使って字幕を付けました。 #00:06:02-6#
- 5 Interviewer: それは直接音声認識でしたか。 #00:06:06-2#
- 6 NHK1: ええとね、ダイレクト #00:06:10-2#
- 7 Interviewer: ああ、ダイレクト #00:06:10-2#
- 8 NHK1: ではないと、リスピーク #00:06:13-0#
- 9 Interviewer: リスピーク、オケー。 #00:06:13-0#
- 10 NHK1: それで、その後ですね、まず大きな方案を説明しますと、字幕はニュース字幕と、あ
- 11 と番組字幕、この二つで分けてですね。生放送に字幕の付与しております。付与して、わかり
- 12 ますか。 #00:06:44-0#
- 13 Interviewer: うん、付与がわかります、字幕を付与する。 #00:06:49-3#
- 14 NHK1: はい、そうなんです。ニュース字幕は、まあ、ニュース決まった時間にあるニュース
- 15 ですね。七時とか、九時のニュースだとかですね、NHK は大体一時間二回ぐらい五分ぐらい

16 のニュースが入ったりするんですけれども、それを対応するチームがあります。はい、あと番
17 組字幕というのは、トーク番組だったりとか、スポーツだとか、そういう物に字幕を付けてい
18 くチームがあります。 #00:07:38-8#

19 Interviewer: ああ、それは分けてあります。 #00:07:41-0#

20 NHK1: 分けてます、はい。で、ニュース番組はまあ。徐々に徐々に、ニュース7からニュー
21 ス9を付けたりとか、そういう形でニュース番組を全般をこう付けていくように徐々に増やし
22 ています。で、すべては音声認識というわけではなくて、特集キーボード、ステノワードでご
23 存じですか。 #00:08:12-3#

24 Interviewer: ああ、分かります。なんか、もっと早くに入れることができます。 #00:08:17-7#

25 NHK1: 高速入力のできる会社に依頼をして付けてもらったりです。 #00:08:27-1#

26 Interviewer: ああ、そうですか。それは他の会社ですか。 #00:08:29-5#

27 NHK1: そうです、あの、委託先があります。外部会社です。協力会社ですね。はい、では、
28 生字幕放送のほうは、音声認識を使うこととあとキーボードでの連携入力をして、字幕を作っ
29 てます。 #00:08:56-6#

30 Interviewer: 分かりました。 #00:08:57-5#

31 NHK1: はい。それで、ですね。2011年に大きいな地震がありました。 #00:09:09-2#

32 Interviewer: そうですね。 #00:09:10-3#

33 NHK1: はい、それをそのあとに生字幕放送への付与率をどんどん上がっているという感じに
34 なってます。なので、2000年からは、少しいつづ付与する番組増えてるんですけれども、
35 まあ、2011年かな、どんどんと番組数が増えていました。 #00:09:44-1#

36 Interviewer: そうですね。現在では、何パーセントが字幕放送ですか。 #00:09:49-1#

37 NHK1: 今は、ええとね、広報テレビは98ぐらい出っただと思うんですが、ちょっとここ、確
38 かな数字をごめんなさい確認してないので、確認して、ご連絡します。 #00:10:10-6#

39 Interviewer: でも、これは全部の放送ですか。生放送だけじゃありません。 #00:10:14-7#

40 NHK1: ええと、全部ですね。 #00:10:19-0#

41 Interviewer: 生放送についての数がありますか。 #00:10:20-7#

42 NHK1: 生放送についてはもう、ほぼついてますね。 #00:10:25-9#

43 Interviewer: ああそうか。 #00:10:25-5#

44 NHK1: うん、日本では朝の七時から夜、深夜12時までの間字幕を付けなさいというふうに、
45 総務省っていうまあ放送局を間隔している、国のなんていうんだ。時間が決まってるんです
46 ね。 #00:10:54-4#

47 Interviewer: ああそうです。素敵ですよ。オーストリアは違うと思います。 #00:10:58-4#

48 NHK1: ああ、違うんですか。 #00:10:59-8#

49 Interviewer: うん、そんなになんか、確かにあの字幕がある生放送がありますけどそんなにす
50 べてじゃないです。 #00:11:07-9#

51 NHK1: ああ、だから、七時から朝の七時から、深夜十二時までの間の番組は今付与するのは
52 こんなふうな番組以外は付けなさいといわれておりまして。現在ほぼついております。ただ、
53 難しい番組があるですね。例えば、日曜討論という番組だって、専門家がスタジオに集まって、
54 一つのことをこうバーン話して行く番組があるのですが、こういうのは中々難しいかなあ。字幕
55 を付けてない。あと国会、政治の国家中継ってもあるですね。はい。これも一部字幕を付与始
56 めましたが、ほぼついてないです。そういう状況なのです。それ以外はついている。それは今
57 の状況です。 #00:12:19-0#

58 Interviewer: うん。では、たぶん、その音声認識のプログラムについてちょっと説明してくだ
59 さい。 #00:12:38-9#

60 NHK1: はい、音声認識はですね、NHK 一軒、技術研究所という場所がありまして、その、
61 そこで開発したものを使ってます。 #00:12:56-9#

62 Interviewer: すみません、それはどこでしたか。 #00:12:58-4#

63 NHK1: NHK 技術研究所です。 #00:13:02-4#

64 Interviewer: 技術研究所。 #00:13:04-2#

65 NHK1: はい、NHK ぎっけんと呼んでます。 #00:13:06-1#

66 Interviewer: うん。 #00:13:17-4#

67 NHK1: 2000年に開発したものをいろいろこう、なんというんでしょう。大切に育ってな
68 がら、今も使ってます。基本設計は変わっていないというじょうきょうです。 #00:13:29-8#

69 Interviewer: ああそうですか。今も同じです。 #00:13:32-3#

70 NHK1: はい。ただね、今年20年間新しい技術が開発されているんですけども、まだです
71 ね。運用に、運用段階まできていないという状況です。 #00:13:52-9#

72 Interviewer: うん、まだ開発中ですか。 #00:13:56-2#

73 NHK1: はい、そういうふうです。あの、開発中ということでもいいです。それで、NHK の音声
74 認識は言語モデルと音響モデルこの二つを辞書をつかって認識を高めています。で、言語モデ

75 ルってというのは。例えば、料理番組でしたら、ハンバーグの話だったら、どんな材料がいると
76 か、そういうキーワードですね、覚え込ます。そういう辞書をつくります。例えば、焼く
77 とか、そういう言葉を辞書に覚え込ませてないとはいけない。まあ、基本的なものは入っているの
78 で、ちょっと例が悪かったんですが、私は車が好き、車のほうにもっていきますと、車の用語
79 は大体はいつてと思えますけれども辞書に、ただテスラって言われた時にテスラってどういう
80 表記でだしたらいいんだろうと。あの、アルファベットなのかカタカナなのか、とか、そう
81 いったことを辞書に登録しなきゃいけないです。 #00:15:18-7#

82 Interviewer: それは、なんか、放送の前に辞書に入りますか。 #00:15:22-2#

83 NHK1: そうです、はい。単語の登録がある、言語モデルを作るという形でやっております。
84 #00:15:32-2#

85 Interviewer: それも字幕キャスターの仕事ですか。 #00:15:33-7#

86 NHK1: それはですね、キャスターではなくて、リスピークした後に、日本語は難しいて、同
87 音異義語とかですね、あの認識しなかった場所もふくめて、修正者がいます。はい、その修
88 正者が直します。なので、その修正者が辞書を作ります。 #00:16:01-7#

89 Interviewer: そうですか。 #00:16:03-0#

90 NHK1: そうしましたら、リスピークの、まず、構成かの話しましょうか。まずですね、われ
91 われ NHK グローバルメディアサービスの人間には ディレクター として、字幕放送今日、どう
92 いう、例えばですね、朝 11 時から一時間の番組があります。字幕を付けることになりまし
93 た。そうしたらこの一時間どういう字幕を番組なのかというも、まず、台本とか構成表という
94 もの参考にしまして、情報を使っています。それを、字幕を修正する人たちに知らせまして、
95 こういう言葉ですね、辞書に入れてねという指示を出します。はい、そうすると、修正する校
96 正者がそれを辞書登録してくれます。 #00:17:09-2#

97 Interviewer: じゃ、その修正者は自分で言葉を選ぶことがしません。それは他の人の責任にな
98 ります。 #00:17:20-5#

99 NHK1: ええとね、ダイレクトの責任になりますので、修正者はあの時分たちでいろいろ辞書
100 を作るために情報を集めてくれたりしますけれども、迷ったら必ず、ダイレクタにこの言葉は
101 どうやって、辞書登録しましょうかとひらがなにしましょうか、カタカナにしましょうか、そ
102 ういうことを聞いてきます。確認してきます。で、リスピークはその場にはいないです。 #00:17:55-
103 0#

104 Interviewer: そうですか。 #00:17:54-5#

105 NHK 1: そうなんです。リスピークはですね、放送が始まる、どれぐらい前ですかね。一時間
106 ぐらい前です、その担当する番組の台本とかを読み込んで、それで番組に読みます。そのさえ
107 に、例えば、この言葉出にくいから、出やすく辞書登録して便利みたいな。そんなやり取りが
108 あります。例えばですね、長いタイトルがあるとします。なんというですか。コロナの
109 [unverständlich, redet im Hintergrund] コロナウイルスってずっとというのは大変なので、タイトル

110 コロナって言ったら、コロナウイルスが出るようなそういう仕掛け、辞書作りの中に出てい
111 きます。 #00:19:09-1#

112 Interviewer: では、それは、その修正者がリスピーカーと務めていないといいますね。 #00:19:18-
113 0#

114 NHK 1: うん、リスピーカーはね、台本読んだ中で出にくそうなものをわれわれそういう登録
115 をですね、タイトル読むというふうです。ですけども、タイトルコロナって言ったら定式
116 のコロナウイルスが出る。そういうことを例えばリスピーカーが指示をしたりこともあちま
117 す。 #00:19:41-4#

118 Interviewer: ああ、そうですか。 #00:19:44-7#

119 NHK 1: でも、基本的にはダイレクタがそういう内職も含めて用語集めて修正を担当する人に
120 辞書登録をお願いするということです。で、NHK では、番組の構成表とか台本のはしっかり
121 できている。そして、リハーサルをします。放送の前。で、そのリハーサルもちゃんとリスピ
122 ーカーがリスピークをしてみて、出るかな出ないかな、出につらい言葉は出やすくするはどう
123 したらいいかなっていうこと放送前にやります。 #00:20:35-9#

124 Interviewer: そして、番組中はリスピーカーも修正者もいますね。 #00:20:46-4#

125 NHK 1: います。ただですね、リハーサルをリスピーカーが参加しない場合もあるんです。
126 #00:20:52-9#

127 Interviewer: ああそうですか。 #00:20:53-7#

128 NHK 1: そういう時はね、修正者がリスピークをやって、リハーサルで一回読んで、読んでみ
129 て出にくい言葉があれば、出やすいようにチューニングをする。そういう作業をします。で、
130 リスピーカーが参加しない番組っていうのは、例えば、前日にリハーサルがあるとか。そうい
131 う番組は、まあ特集ですけども、参加しない場合があります。 #00:21:23-2#

132 Interviewer: そうですか。 #00:21:27-7#

133 NHK 1: でも、ほとんどね、リスピークの番組にはちゃんとリハーサルが当日行われますで、
134 リスピーカーがちゃんとリスピークをしてみたらというの、出るか出ないかというリハーサ
135 ルでほぼやっております。で、あと、それは言語モデルです。で、音響モデルというもありま
136 す。それは、辞書は言語モデル、音響モデル二つで作ってあるんですけども、音響モデルは
137 話すリスピーカーの声を覚え込ませているものです。例えば、ヨハンナさんの声と、認識結果
138 が落ちするように、あの、例えば、ヨハンナさんがバーと話した言葉がですね、ちゃんと音声
139 認識で、正解してないところを修正して全部正解にしたりいる、その音声のデータと正解デー
140 タ、変更付けてあげているものが音響モデルです。 #00:22:47-2#

141 Interviewer: そうですか、それは各リスピーカーのモデルということですか。 #00:22:52-5#

142 NHK 1: そうです、はい。各リスピーカーのモデルです。 #00:22:55-7#

143 Interviewer: それはオーストリアでもそうです。 #00:22:58-3#

144 NHK 1: ああそうですか、じゃ #00:23:00-2#

145 Interviewer: 他のプログラムですけど、同じです。 #00:23:05-4#

146 NHK 1: オーストリアはちなみ、どんな仕組みでリスピークしてるんですか。なんか、オリジ
147 ナルのソフトですか。 #00:23:11-5#

148 Interviewer: ううん、オリジナルじゃない、あの、DRAGON NATURALLY SPEAKING と
149 いう ソフトです。 #00:23:16-3#

150 NHK 1: ああ、ドラゴンスピーチ。 #00:23:18-6#

151 Interviewer: そうです。 #00:23:18-9#

152 NHK 1: はいはい。それは、日本でもありますね。 #00:23:22-9#

153 Interviewer: ああ、ありますか。 #00:23:23-8#

154 NHK 1: うん、あります。わかりました。 #00:23:30-6#

155 Interviewer: そして、リスピーカー、修正者がいません、リスピーカーは自分で修正をします。
156 #00:23:37-0#

157 NHK 1: リスピーカーが修正 #00:23:37-9#

158 Interviewer: そうです。 #00:23:38-0#

159 NHK 1: すごいですね。 #00:23:39-5#

160 Interviewer: はい。 #00:23:40-6#

161 NHK 1: それはね、日本語はご存知の通り、それは無理ですね。 #00:23:44-9#

162 Interviewer: 確かにそうです。 #00:23:46-4#

163 NHK 1: 同音異義語があるし、 #00:23:47-5#

164 Interviewer: 漢字も、うん。 #00:23:49-2#

165 NHK 1: 今ですね、リスピーカー一人にが話した言葉を修正者二人がああ修正してます。
166 #00:24:00-3#

167 Interviewer: 二人ですか。 #00:24:00-8#

168 NHK 1: そうです、はい。二人が修正してます。リスピーカーがしゃべった言葉を二人の修正
169 者が、一つの文章を半分を A さんが、半分を B さんが、っていう形で修正してます。今まで
170 は、リスピーカー一人、修正者一人で、放送に、放送担当してたんですけども、修正者の負担
171 はすごく多かったの、大きかったの、修正者二人にしたんです。それが始まったのが、
172 いつでしたけ、それは、ちょっとまってください、聞いてきます。 #00:25:05-4#

173 Interviewer: 大丈夫です。 #00:25:13-1#

174 NHK 1: 2 0 1 4 年か、2 0 1 5 年かそうです。ダブル修正したのは。 #00:25:40-9#

175 Interviewer: そうですか、2 0 1 5 年でした。 #00:25:45-5#

176 NHK 1: うん、2 0 1 4 年か2 0 1 5 年とこっちかわかないので、なの。 #00:25:50-1#

177 Interviewer: オッケー #00:25:50-9#

178 NHK 1: 今は、2 0 2 0 年。だから、五年前に。 #00:25:57-5#

179 Interviewer: 大丈夫です。 #00:25:59-2#

180 NHK 1: 大丈夫ですか。ダブル修正にしました。で、なのでリスピーカー一人修正者二人って
181 いうのが、基本ですが、短い番組ならで、リスピーカー一人でいいんです。長い番組になる
182 と、二人になります。大体1 5 分となると二人になります。はい、1 5 分になったぐらい、リ
183 スピーカーが変わりますので。 #00:26:28-8#

184 Interviewer: 1 5 分、はい。 #00:26:28-8#

185 NHK 1: はい、なんか同時通訳をすると同じぐらい疲れるらしくで、 #00:26:36-8#

186 Interviewer: 2 0 分だと思います。オーストリアでは2 0 分です。 #00:26:35-6#

187 NHK 1: そうですね、今ね1 5 分ぐらいで変わってます。 #00:26:50-2#

188 Interviewer: 修正者も変わりますか。 #00:26:53-3#

189 NHK 1: 修正者は、修正者はですね、だいたいリスピーカーが二人対戦時は修正者三人います。
190 #00:27:01-5#

191 Interviewer: そうですか。 #00:27:02-6#

192 NHK 1: それで、あの、??ところでかわります。同じくらいかなあ。そうですね、ただですね、
193 変わっても休んでるわけではなくて、ちゃんと放送を聞きながら、ただしくリスピークをして
194 るかな、修正してるかというのは、その人もチェックしてます。 #00:27:28-6#

195 Interviewer: うん、もちろんそうです。 #00:27:30-0#

196 NHK 1: わずが、働きますね。はい。で、生字幕は日本語から日本語だけですといますのは、
197 はい。他言語から日本に翻訳するていうことはまだやってないですね。 #00:27:58-3#

198 Interviewer: それなんか、計画とかありますか。 #00:28:00-9#

199 NHK 1: 一応いいてのか、あるのかわかなくてですけど、オリンピック、パラリンピックで、英
200 語から他言語っていうのやってみようかって言うことで、してたんですけど、うん変気なっち
201 ゃいましたね。 #00:28:21-1#

202 Interviewer: そうですね。 #00:28:22-4#

203 NHK1: これは、放送ではなく、ネットでやろうでいったんので、放送ではまだないです。
204 他言語かっていうのは。ただ、あの私たちはNHKの子会社なのでNHKがどうするかという
205 ことに関してはあのわからないんですよ。なので、将来的にそうするかどうかというNHKの
206 展開まではちょっとわからないです。 #00:29:00-3#

207 Interviewer: そうですか。 #00:29:04-6#

208 NHK1: はい、NHKに聞かないといけないですね。 #00:29:09-3#

209 Interviewer: はい。 #00:29:12-5#

210 NHK1: ただ、公式な見解が必要ですね。そうすると、他言語に関してそうすると、なかなか
211 NHKが答えるかどうか、っていうのはちょっとわからないですね。 #00:29:34-4#

212 Interviewer: はい、わかりました。 #00:29:36-7#

213 NHK1: はい、もしあれだったらNHKのホームページから、日本語でも英語でもいいので、こ
214 ういう見解があるかどうか答えてくださいって連絡すると返事があるかもしれないね。 #00:29:55-
215 2#

216 Interviewer: はい。 #00:29:57-7# #00:29:57-4#

217 NHK1: あの、例えば先おっしゃてた、現在で生放送何パーセント出てるんのかとか、それは
218 NHKの編成というところがデータを持ってるですけども、くれるデータがあれば答えてくれ
219 ると思います。 #00:30:17-5#

220 Interviewer: はい、ありがとうございます。 #00:30:17-5#

221 NHK1: NHKのホームページだ。お問い合わせみたいもあるか、コールもあるみたいです。そ
222 うですね、そこでなか、電話じゃないですが、メールで聞いてみるというはあるかもしれな
223 い。そしたらそれが公式見解になると思います。はい。で、今校正者全員がリスピーカーの仕
224 事をしてますかという、リハーサルではします。 #00:30:43-9#

225 Interviewer: リハーサルで。 #00:30:45-6#

226 NHK1: 校正者が、リハーサルでリスピークをするっていうことがある。で、リスピーカーが
227 校正するっていうことは今ないですね。 #00:30:56-1#

228 Interviewer: ないです。 #00:30:56-0#

229 NHK1: はい、ないです。で、放送中は、リスピーカーは変えますけども、リスピーカーと修
230 正者が交代することがないね。リスピーカーはリスピーカーの仕事するです。交代は先に言っ
231 た、まあ、15分間かくぐらいて行われていますが、休憩というよりはも、あのお互いチェッ
232 クしているので、放送中休憩はないですね。で、新し人員を探す時重視している能力や資格と
233 いうところでも、あの、私たちがNHKグローバルメディアサービスが、リスピーカーや修正者
234 のこう就活している会社をお願いをして業務お願い受けてもらっている感じです。それで、業

235 務委託という形で出してますんで、その人員がですね、その仕事ができるっていうことは大選
236 定なので、そうですね。 #00:32:13-1#

237 Interviewer: まあ、でも、多分須藤さんの意見によると何が重視ですか。 #00:32:18-5#

238 NHK1: まず、リスピーカーはほぼ、もっとアナウンサーとが、キャスターです。 #00:32:26-1#

239 Interviewer: そうですね。 #00:32:27-3#

240 NHK1: はい、それは、ほとんどは頭でまとめていく力がある人とあと活舌がいい人。こうい
241 う人がリスピーカーをやってます。修正者は、どうですかね、メンタルの強い人ですね。放送
242 の中でも、けっこう慌てずとし、的確に崇敬をしていく人。だから、適正というもあると思い
243 ますね。だれでもできる仕事ではないと思います。 #00:33:10-4#

244 Interviewer: もちろんそうです、うん、 #00:33:12-8#

245 NHK1: ね、そう、特に日本語を知らないと、日本語は難しいですよ。こえるっていう漢字が
246 ね、いくつもあるわけですし。どっちの、どの漢字を使ったらいいかなあというも順次に判断
247 しなきゃいけない。 #00:33:32-1#

248 Interviewer: そうですね、私ぜんぜんできません。 #00:33:36-4#

249 NHK1: その判断するために、NHK は用字用語辞書というのがある。その辞書を読むと答えが
250 書いてあるんですけども、生字幕中で辞書引けないので、みんなさん大体頭に入って、辞書
251 で結構厚いんですけど。はい。結構がついて、日本語辞書丸々覚えているような感じですね。
252 #00:34:07-4#

253 Interviewer: すごいです。 #00:34:08-2#

254 NHK1: だって、構成表が台本があるので、どんな言葉が出るかなというのはだいたい分かる
255 ので、以前に辞書をひいて、こういうが出てきたらその漢字使うとか、その字を付かなきゃい
256 けないだの。そういうのは当たりあうを付けています。はい。 #00:34:39-2#

257 Interviewer: はい。 #00:34:40-0#

258 NHK1: 字幕の品質はどのようにチェックするかというと。ガイドラインね、これはあるかない
259 かって、ガイドラインがないと言ったら多分、それはヨハンナさんが書いた文にでたら中々
260 あるかないか僕は見たことないけど、多分でもある。それはNHKのほうに聞いてもらったほ
261 うがいいですね。そのガイドラインにはありますかと、品質のチェックっていうのはですね。
262 編成担当の、まあ、チェック担当の人間を別に放送の時に立ち会えをさせてます。それで、そ
263 の人はずっと音声を聞きながら出てる字幕が正しいかどうか、漢字が間違ってたかな、
264 というのタイプで確認をする人はいます。 #00:35:46-4#

265 Interviewer: そうですね。 #00:35:47-2#

266 NHK1: はい、大きなミスはなかったかなとかね。でも、ミスがあればこういうミスがあった
267 というのを書き留めて、込みます。ただね、生字幕でこんなミスしたんだって、とんでもない

268 なあっていうふうに視聴者からクレームを貰うことがあまりめったにないですね。 #00:36:08-
269 4#

270 Interviewer: そうですか。 #00:36:11-6#

271 NHK1: そうです、はい。では、放送法では十分間字幕が出なかったら事故です。オーストリ
272 アではありますか。放送事故。何分出なかったら字幕がでなかったら。放送事故。 #00:36:29-
273 5#

274 [...]

275 NHK1: 日本は十分出なかったら放送事故なんです。だから、その字幕の品質っていうに関し
276 ては、その品質だけをチェックする人間が放送に立ち会ってます。 #00:37:12-2#

277 Interviewer: そうですか、でも、適当なフォーミュラとかチェックするためにいろいろなフォ
278 ーミュラがあります、例えばNERモデルということはありませんか。それはありますか、こうい
279 うことがありますか。どういえ。 #00:37:52-1#

280 NHK2: 例えば。。。 #00:37:50-5#

281 Interviewer: 例えば、あの、このテキストはこの字幕では何語は誤りがありますかという計算
282 とかありますか。そんなことをしませんか。 #00:38:12-6#

283 NHK1: ああ、例えば何十分の番組でいくつ間違えた単語がありますか。 #00:38:17-3#

284 Interviewer: そうです。 #00:38:17-3#

285 NHK1: ああ、それはね一応チェックしてますね。リアルタイムで。どこ間違えたというの
286 [undeutlich] する人がいて、その人があの時この字間違えましたねとかそういうのはチェック
287 して。 #00:38:36-0#

288 Interviewer: そうですか。 #00:38:38-2#

289 NHK1: はい。 #00:38:42-0#

290 NHK2: もしかしたら、日本特集かもしれないんですけど、世界中そうするかもしれないです
291 が、NHKは放送する前にあの、今から、生字幕ですよ、間違えるかもしれませんよっていう放
292 送する前に立ちます。TVプログラムの前にあの、今からは、できる限りやりますけど、間違
293 えるかもしれませんっていうのは。決まった文章を対象出してるんですよ。生放送のため、
294 字幕で表現できない場合があります見たいな、そういう文章ですね。まず、お知らせしてきま
295 す。だから、まちがっても許してね。そういうのをいれてますね。いれてます、オースト
296 リアでは。 #00:39:41-9#

297 Interviewer: ないです。 #00:39:43-9#

298 NHK1: ないですね。はあ。 #00:39:47-3#

299 Interviewer: まあでも、別にクレームとかがないとおもうけど。 #00:39:54-0#

300 NHK1: いい国ですね。日本ではそんなにクレームがないですよ。でも、人の名前を間違えっ
301 たりとか、あの地域の街が絵ったりするとそこに住んでいる人たちはあの、だれが分かった
302 りする人たちかなあ。こう違う、違ってよ、という指摘はいただいたりしますね。だから、台
303 本にあることはね、事前準備して間違えないんですけど。放送って、生放送の場合何が出るか
304 わからない番組もあるんですね。突然話した盛り上がりちなって、台本もないことしゃべり始め
305 っちゃ。そうすると言語モデルがないので、出ないんですね。その出ない言葉が例えば昔の映
306 画の名前とか、昔の俳優の名前とか、私たちが若いのでおじいちゃん、おばあちゃんが好き
307 だった俳優の名前なんて現れてもわかりません。そういうの間違いがあると、間違えってるよ
308 てでなくなった。 #00:41:12-2#

309 Interviewer: あの、その時はどうしますか。 #00:41:13-2#

310 NHK1: ひらがなで出します。 #00:41:19-7#

311 Interviewer: ああ、ひらがなで、オッケ。 #00:41:20-4#

312 NHK1: はい、ひらがなで。ただ外国の関係するなら、カタカナです。ただ、漢字とかだとね、
313 間違ったりするので分からないものはひらがなででってするというにします。 #00:41:41-6#

314 Interviewer: はい。 #00:41:47-5#

315 NHK1: そうすると、形式について、話者の交代は、話者交代マークっていうの使ってるんで
316 すけど。それはどうしたらいいかな。… #00:42:08-5#

317 NHK2: ああ。リスピークするの、交代。 #00:42:13-1#

318 NHK1: 話者の交代でリスピークではなくて放送の時に人が変わったよっていることどう見せ
319 るかということか。 #00:42:19-5#

320 Interviewer: ああ、はい、そういうことです。 #00:42:23-4#

321 NHK1: どうしよう。じゃ…カメラをにします。これでいいかな…見えるかな。 #00:42:43-2#

322

323 Interviewpartner zeigt dieses Symbol: << #00:42:51-0#

324 NHK1: ええとね、見えるかな…こんな、そんな感じ。 #00:42:58-6#

325 Interviewer: ああ、分かりました。こういう。 #00:43:04-6#

326 NHK1: ジョンジョンっていうこうなんか、こんなマーク。それを二つつけ、そんな感じで、
327 それ見えるかなあ。ここで、井上でこれ話者変更マーク。 #00:43:29-5#

328 Interviewer: ああそう、そうわかります。 #00:43:38-5#

329 NHK1: コロナために今マスクをしています。そういう漢字でございます。 #00:43:46-8#

330 …irrelevant #00:44:35-7#

331 NHK1: 後はね、句読点はね、使ってますます。でもこれは、あの、音声認識で話と自動的に
332 句読点を付けます。 #00:44:45-6#

333 Interviewer: ああ、自動的ですか。 #00:44:48-1#

334 NHK1: うん、足りないところを、修正者が入れています、ここに句読点があったほうが、わか
335 りやすいねっていう時はあの修正者が入れています。で、改行突端も修正者がしています。で、生
336 放送は今の15.55字一行15.55字。それが二行で放送しています。一行15.55字書
337 ける二行で放送します。そこにね、話者があれば、話者変更マークも入れていますね。はい、で、
338 気を付けてるのはあの通じがこう、途中に書き終われちゃうと結果が変わって意味になるって
339 そうならないように、気を付けて鍵です。 #00:45:47-2#

340 NHK1: で、字幕の遅延について、遅延は[unverständlich]長いに違くないですね。リスピークも認
341 識が悪い場合は、本当に遅れると30秒遅れてるもありますね。 #00:46:04-6#

342 Interviewer: 30秒。 #00:46:05-3#

343 NHK1: 早いともう六秒でできます。リスピークして、修正してから放送まで六秒くらいかな。
344 でもそういうともう中ほんとに30秒かかっちゃう場合もある。 #00:46:23-8#

345 NHK1: そうですか。 #00:46:25-1#

346 NHK1: そうするとね、非常にづらいので、そういう時は溢れてるなという時はリスピーカー
347 が話す分量を一致に減らします。それで、表示リアルにあわせる。で、そういうのはそうしま
348 す。でも、そんなに遅れるのはめったにないですね。許容範囲にはまあこれは中々NHKに聞
349 かないといけません。 #00:47:05-3#

350 Interviewer: そうですか、オッケ。 #00:47:05-3#

351 NHK1: 僕らが答えた時間問題になります。これぐらいつかってというと。ただ、見ててわかる
352 というのは一番重要です。絵と音声が多に離れてしまったらわからないと意味ないの
353 で、だからリスピークの量を調整しながら、字幕を極力にあわせるようにやっていますね。で、
354 番組の放送意図的に遅らせてということは、一回だけあのたったことがあります。 #00:47:46-
355 3#

356 Interviewer: 一回、オッケ。 #00:47:48-2#

357 NHK1: はい。そういうことができますねという確認はできてるので。ピッタリ字幕って私の
358 読んでますけどもピッタリ字幕一回だけやっています。それは放送30秒遅らせてやりました
359 ね。スタジオの映像を見ながら、それはね、でも、リスピークじゃなくてキーボードでやった
360 んです。 #00:48:21-4#

361 Interviewer: ああ、どういう番組でしたか。 #00:48:22-0#

362 NHK1: ええとね、総会社の方たち為の番組。バリバラというばんぐみでやりました。それは
363 あれでした。間違ったこと言った異変ないです。ヨハンナさんこれ発表する前に一回長大あ
364 の、こういうに答えてくれてますよね。ちゃんと了承取る、そのほうが安心で...#00:48:45-5#

365 #00:48:56-1#

366 Interviewer: いいです。 #00:48:55-4#

367 NHK1: 大丈夫ですか。まじめですから。。。 #00:49:03-5#

368 Interviewer: いや大丈夫です。別に。。。 #00:49:11-1#

369

370 NHK1: 字幕がピッタリです、一回だけ。去年。去年やってました。でも、これをやらせれ
371 てにはNHKの編成だったり、番組を制作するチームだったりあと字幕のチームだったりこれ
372 をこんな三つかですね、一緒に動かなきゃいけないです。中々まま、実現を難しいですけど。
373 大体キーボードだとはつから15秒でちゃんと成型していけるですよ。ただね、リスピークは
374 修正に時間かかるとすご30秒とかなっちゃうんで、なかなかピッタリ字幕やるんだったら一
375 分ぐらいディレイもってますとかもしれないです。キーボードなら30秒でできるけど。もと
376 の言葉通りを字幕を作ることはどの程度基本的にはですね、あの、もとの、初は放送で話して
377 る通りに字幕を作ることは目指してます。つかっていうとあのこの後にも質問がありましたけ
378 ど、聴覚障碍の人たちとこう向いたお願いしてましたどういう字幕が見やすいとかということ
379 はいつも見せるんですけれども二つのタイプがいるんですね、省略してもいいかな、映像にで
380 きるだけスピードあげてくれて、もう一つのタイプは、スピードあげてくれてもいいかな何言
381 いてるか全文教えててくれ。そういう二つのタイプがあります。なので、できる限る言葉通り
382 に字幕を作るんですけれども、遅延、遅れが大きくなってきたら先を話した通りテキストを減ら
383 してできるだけスピード調子にする。追いついて来たなか話者が増やしてとかそういうなこと
384 をして。あとですね、日本語テレビで結構。。。多いんです。 #00:51:42-6#

385 Interviewer: 何が多いんですか。 #00:51:46-6#

386 NHK1: 我々はクローズドキャプションで字幕を出します。それに対するオープンキャプショ
387 ンって言ってテレビの中にともどんどんコメントフォローをしてるんですね。そんなのも、見
388 ればわかるからそういうもの理解してはリスピークをしないでコメントを見せる。それによっ
389 て、ちょっと字幕の遅れって取り戻すね。そういうことにします。で、もっというとスタジオ
390 の部分とVTRの部分というのはありますね。VTRはもう事前にできてるものなんて、番組
391 中もらってコメント全部も出し残した。だから、スタジオ部分だけリスピークしてVTRにな
392 ったらリスピークなしにします。 #00:52:39-1#

393 Interviewer: ああ、オッケ。 #00:52:38-1#

394 NHK1: それで、あのリスピークの遅れはそこで取り戻せたりすることですね。はい。わかっ
395 てくれたりですね。 #00:52:52-8#

396 -- #00:53:08-8#

397 NHK1: 22 番が、生字幕を充実させる予定はありますか。充実はなんか今 NHK は双方テレビ、
398 教育テレビにですはニュープレミアムだと四系、[unverständlich]それで字幕を一生懸命できる
399 限り多くつけて行きましょう。それ以外で他の時間に拡大するというの充実させていくという
400 はありますね。あとはクオリティもよくすると遅れの無い、間違えがない字幕をやっている
401 ということで、みんな毎日汗をかいてやっております。で、あとはまあ、生字幕の未来につい
402 てやっぱり音声認識を充実するところによって、できる限り時間が増えて行くんじゃないかなと
403 思っているんですけど多分日本語が世界一難しいと言ったら送られちゃうかもしれませんが、
404 漢字が、ひらがなが、カタカナがあって、かつ同音異義語がある、同音異義語がわかります。
405 #00:54:23-7#

406 Interviewer: ああ、わかります。 #00:54:25-7#

407 NHK1: うん、同じ漢字 #00:54:30-9#

408 Interviewer: 意味が変わります。 #00:54:34-9#

409 NHK1: それをこうやって行くにはちょっと音声認識は結構大変です。これはわれわれの中で
410 の認識ですけど、音声認識が 95 パーセント以上を認識しないともね、キーボードの連携入力
411 のも早いです。だから日本語 95 パーセント以上を認識するように辞書を使ったり音響モデル
412 を充実させたりしてやっています。ああ、一番なんかあれです、同音異義語が書きありそうだと
413 でした。あのここまでなんかあります、ご質問は? #00:55:22-3#

414 Interviewer: あの、直接認識について、それもしてますか。 #00:55:29-6#

415 NHK1: ダイレクト認識はですね、ニュース字幕では詰めてます。 #00:55:37-2#

416 Interviewer: アナウンサーの声 #00:55:37-9#

417 NHK1: うん、アナウンサーが一人でしゃべるので、そのマイクの音声をダイレクトで認識さ
418 せるっていうことしてます。ただ、ニュース字幕でも街頭インタビューとかそういうところはざ
419 わざわしてるので、認識の落ちるのでそこだけ修正者がリスピークしてます。ニュース字幕
420 の構成は修正者が二人、ダイレクト認識させたの二人のオペレーターが修正します。ただ、三
421 人目がねさまに座ってて認識を落ちてそういった街頭インタビューは、三人目がリスピークし
422 ている。大体ね、五分のニュースで、街頭インタビューが 20 秒に内科ぐらいですけど、ただ
423 ね、街頭インタビューって大体おペンキャプションがはいってる。こういうところもはいてるん
424 ですね、ま、しなくてもいいですか。あとは内緒ですけど、安倍総理大臣の音声は全然ダイ
425 レクト認識してもある。 #00:56:57-6#

426 NHK1: 生字幕未来ってやっぱり音声認識っていうのははいてくるでしょうけど。でも、全部
427 に音声認識に任せるってことは絶対だめなので、それは必ず人間が修正して出していくと
428 いうことは多分変わりがないだろうな。あとは、日本語でテキスト化するわけですから、テキ
429 スト化した日本語今度他言語化していくかね。それで、日本にいるあの外国の方々にあのそれ

430 つなぐ情報を国の言葉で伝えられるになったらいいんじゃないかな。こう個人的にです。

431 #00:59:23-3#

432 Interviewer: 分かります。 #00:59:23-5#

433 NHK1: じゃ、リスピーカーへの質問。いつリスピーカーとして働き始めましたかという、結
434 構ね、リスピーカーさん長い人が多いんですよ。そうここが来る前からリスピーカーさんが
435 いるから、おそらく長い人でも 15 年ぐらいやってる人がいるんじゃないですかね。そう、た
436 だですね、新しい人があまり入ってないです。 #01:05:23-9#

437 Interviewer: そうですか、どうしてですか。 #01:05:24-5#

438 NHK1: あの、多分年に一回ぐらいに NHK のアナウンサーがリスピーカーとして入ってこ
439 とありますね。それは、リスピークで一番長いのが、相撲の放送があるです。相撲の時間は毎
440 日毎日あのリスピークが必要なので、で、相撲はだから、何年間ですか。二千年二年大相撲
441 六日間だけでしたそうです。で、二千三年からもう、大相撲初爆 15 日間 15 日間の従業があ
442 るですけど、その 15 日間全部二千三年からずっと放送してます。ずっとだったらもう、何年
443 か、二千三年だったら。17 年。リスピークでやってるので、辞書がすごいんですよ。 #01:06:46-
444 5#

445 Interviewer: ああ、そうですね。 #01:06:45-8#

446 NHK1: うん、17 年間育ってた辞書だから、なんか 10 年に一度の決まりいてて難しい決ま
447 り、なんか、技が技の名前を詣でるという。で、ここでは毎日毎日リスピークがあるので、
448 人が入る時があります。でも、その人も NHK のアナウンサーで、ええと、お手伝いに来る
449 漢字ですか。 #01:07:28-2#

450 Interviewer: でも、そういうリスピーカーは相撲だけに担当していますか。 #01:07:29-2#

451 NHK1: だいたいね、スポーツ、相撲以外野球もリスピークでやることが、あとサッカーこう
452 というスポーツは長時間なので、リスピークがが三人、四人ぐらいでもね。 #01:07:45-1#

453 NHK2: 四人。 #01:07:45-7#

454 NHK1: ごめんなさい、四人ぐらいでかかもしれません。相撲はね、四人ぐらいがいます。だから、
455 サッカーをやっている人もそのぐらいですね。 #01:08:01-2#

456 Interviewer: でも、その人は相撲だけをリスピークしていますか。例えば野球とサッカー #01:08:10-
457 7#

458 NHK1: 相撲も、野球もやります。 #01:08:13-0#

459 Interviewer: でも、同じ人が、相撲も野球も。。。 #01:08:13-7#

460 NHK1: 同じ人です、はい。まあ、相撲があると野球があることがないですけどね、一日一番
461 組ぐらいですけども、はい。集中力が持ったないですよ。で、字幕キャスターになる以前
462 というのは大体皆さん、どっかの放送局でキャスター経験があるかた、テレビでリポーターを

463 やってたりとか、キャスターやってたり。だからテレビでしゃべったことのない人というのは
464 リスピーカーにはまだいないですね。だからね、コストが高いんです。リスピークのコストた
465 げるにはプロのキャスターとかじゃなくて修正者とかがやっていくになるとコストが下げられ
466 るけど、そういうことをもしくぼくは後ろからはされちゃった去年。内相です。で、リスピ
467 ークのキャスターの教育っていうのは、大体あの先輩のリスピーカーは、教えてます。ただ、
468 それはもほとんどテレビに出てアナウンサー経験がある人はすぐできますね。なぜか知らな
469 いけど、すぐできる。リスピークっていうと同時通訳と同じ公なんて言うんでしょう。聞き試
470 しながら話していくっていう形ですよ、結構大変なんですよ。で、これをこうテレビとか
471 見ながらリスピークしてなれていくみたいな感じだと思いますけど。教育としては、今大体新
472 しくリスピーカーになりましたっていらっしゃたアナウンサーたちもうはじめからできてます
473 ね。どこでどう練習してるのかよくわからないけど家でずっとリスピークしながらじゃないの
474 かなと思うぐらい。 #01:10:54-0#

475 NHK2: もともと実況者ですとから #01:10:54-3#

476 NHK1: ああそうそう。もっともっと相撲とから相撲で実況者たちですから何を番組でいうの
477 かっていうのわかっている。笑い話ですけど字幕お相撲で放送で言ってないのに字幕に先に書
478 き出した。これ今はないですけどね。なんか昔はそうなのかあったというのっていうの。それ
479 はね、後輩が今テレビで読んでたです。昔はテレビで読んで人がリスピーカーをやってるてっ
480 ていう時には、大相撲のコメントはこうやるんだろう見たいな形でした。そう。他はいけない
481 です。放送は見たままやらない。大相撲だけなんか昔そんなことがあったというはきいたこと
482 ある。どんな種類の番組をリスピーカーでやってるかという、複数話者の番組っていうのは
483 結構リスピーカーでやったんです。後はスポーツ。スポーツも実況と解説ってあるですけど実
484 況部分は見てればわかるんです。なので解説部分をあのリスピークしていくっていうな形で見
485 てる人の[unverständlich]を減らしてます。で、複数話者の場合、例えば「朝一」という番組が
486 あるんですけど、これはゲストとか大体5人ぐらいはしゃべるんです。その時に今だれかコメ
487 ント生かしたほうがいいかな、あとかいうのリスピーカーが一人で判断してリスピークして
488 いくのであのわかりやすいですね。複数話者の場合はこうリスピークするのは結構有効です。
489 で、NHKではリスピーカーが色分けましています。司会者はみどり。 #01:13:05-5#

490 NHK2: 司会者は黄色。 #01:13:10-3#

491 NHK1: 黄色。専門家緑。あとは、あと何です。ゲストがしろとか。そうなのリスピーカーが
492 いろわけをしてだして。 #01:13:24-3#

493 Interviewer: そうですか。 #01:13:21-8#

494 NHK1: はい、ただリスピーカー大変なんです、リスピークするだけじゃなくて、色分けボツ
495 タンでもあってそれを一度おしながれ一仕事なって。 #01:13:36-3#

496 Interviewer: ああボタンですか。 #01:13:38-0#

497 NHK1: さっきの話者は、だれだれコメントっていうこと話者マークをいれて色分けをして見
498 やすく出してる。リスピーク方式も毎日やってます。朝の番組で、「朝一」という番組、朝 8
499 時 5 0 分から 1 0 時 5 0 ぐらいです。一時間、二時間、二時間 4 5 分ぐらいかなあ毎朝リスピ
500 ークで放送出してます。そう。生字幕以外に仕事があるかどうかわからないんですけども、
501 一人の方はラジオのキャスターをやってらっしゃる。その方はけっこう有名なラジオキャスター
502 です。あと NHK のアナウンサーが二人とかそういう人たちがいらっしゃるんです。就業時間
503 はも番組によって違うんですが大体放送の一時間ぐらいまえきて放送対応するとだから一時
504 間の番組だったらリハーサルもあるんですよ。リハーサル丸々参加すると、だから就業時間
505 はくらは大体一時間の番組によけるのに。 #01:15:26-0#

506 Interviewer: でも、フルタイムですか。 #01:15:32-3#

507 NHK1: フルタイム、番組が終わったらねも帰ります。 #01:15:42-5#

508 NHK2: フルタイムか、ハーフタイムかということですね。ハーフタイム就業ですね。 #01:15:44-
509 6#

510 Interviewer: そうです。 #01:15:45-2#

511 NHK1: パートタイムな仕事ですね。フリーのキャスターとかが多いので、パートタイムジョ
512 ブですね。ただ、NHK まアナウンサーも来るですね、そういう人たちはなん、でも字幕だけ
513 の仕事じゃないですからね、アナウンサーのための、なんというんだろう。教育をしたりとか
514 そういう人もそういうことやってるので。NHK の職員か職員じゃないかというところもまた
515 ちょっと違うかもしれないね。そう、リスピーカーでは NHK の職員がやる場合と、仕事の一
516 部ですね、あの元職員ある場合と、あとは業務委託で、アルバイト、なんというパートタイム
517 で言ってもらおうという場合もあります。一番よくある問題や間違いややっぱり同音異義語だよ
518 ね、同音異義語はけっこう苦労してますよね。技術的な問題はねそんなによくありません。2 0
519 0 0 年ぐらいからずっとやてるので、技術的な問題はね、ほぼないですね。ただやっぱりあの
520 パソコン使ってるので、修正者はタッチパネルなんです。ただモニターの感度がわるいと
521 か、そういうのはときどきありますね。でも放送は常に予備を用意してるのですぐ切り替えれ
522 るのでそんな大変なことにならない。ストレスはねダイレクターを攻めることですね。リスピ
523 ーカーがダイレクターをもっと一体情報くださいわれわれがストレスはさんでし。大体はね、
524 リスピーカーが女性が多いんです。おぼ女性ですよ、NHK のアナウンサーがリスピーカー
525 になるときはほぼ男性ですけど、[...]の番組はもほぼパートタイムは女性。すごい怖い人た
526 ち #01:18:29-9#

527 Interviewer: 怖い？ どうして？ #01:18:33-3#

528 NHK1: 怖いよね、先輩ですから。仕事に対してすごい真面目でね。やっぱりいい字幕を出し
529 たいっておもいは強いですね。だから、われわれに厳しくあったってくる。僕たちがストレス
530 高いです。半分冗談が半分本気です。そう、字幕を付与する際の一番難しい状況はどういう状
531 況ですか。一番難しい状況は番組の資料がない時であるんですよ。そういう時はやっぱりリス
532 ピーカーさんすごいあの大変そうですね。例えばあの一方の番組の中でええ、中継のコーナー

533 があります。でも、中継の場所がわかってるけれども、どんなこと中継するのかわからない。
534 そういう時があったんです。それはね、怖がってますよね。僕たちも怖がってますけど、リス
535 ピーカーさんは怖がってますよ、本当に。そういう場合はインターネットでその場所とかいろい
536 ろ調べてええと、こんなこと中継するんじゃないかなということで、われわれいろいろ情報集
537 めてリスピーカーさんに渡ししたり。それが外れるとリスピーカーさんもっと怖って
538 ない。で、同音異義語とかとかいてあった、同じ読み方の漢字何割ぐらいはシステムが自動的
539 に修正します。システムは自動修正してくれないので、なので多すぎがった時に校正者がもう
540 あのすぐ修正します。ただ、われわれのシステムは同音異義語があったら変換方法っていうの
541 はバット出ってくる。だから、選択すればすぐいけられるので。ただ、同音異義語の場合ね、
542 やっぱり日本人として間違得たくないあのところがあるね、結構校正者の国語力が壊れます
543 ね。ただあの責任はダイレクターの責任になります。どんな間違えが出てもダイレクターの
544 責任になります。校正者とかリスピーカーへの責任に渡してるっていうことはないです。謝
545 るのはダイレクターで、… #01:21:27-0#

546 Interviewer: それは、日本ばいですね。 #01:21:28-3#

547 NHK1: はい、われわれはごめんなさいをするっていう感じ。字幕の何割ぐらいはもとの言葉
548 通り。元の[unverständlich]だろうね。言葉通りっていう余計なところは全部けっちゃいますよ
549 ね。そういうのは全部けっちゃうので、これはあの元の言葉通り逃げるんなあ自体出ないも
550 ん。そういうを考えるともうほぼ言葉通りです。ただやっぱり言葉を丸めてるというですね。
551 あの、分かりやすく、言い換えるっていうことがあります。でもどれぐらい言い換えてるか
552 いうと番組によってやっぱり違ってきちゃうです。番組っていうか、誰が話してるか?? 話し
553 ている人は分かりやすい言葉に言い換えてすることがあります。 #01:22:45-4#

554 NHK1: リスピーカーにとって一番大切な資格…ああ…それはリスピーカーからしてみるとや
555 っぱり耳がいいというところじゃないですかね、聞き取れる。なんかある、語尾や弱い人とかも
556 いるけれども、ちゃんと聞き取れるというのはしゃべるっていうこともそうですけどちゃんと
557 聞き取れる、単語を聞きなれない単語ももちろんリスピークできるっていうこと。そういうと
558 ころはあのリスピーカーにとって必要。耳がいいか悪いかそういうところかもしれない。た
559 とえ[unverständlich]それはいい耳を持ってないとはいけないかな。これはなあ個人的に思いま
560 すね。 #01:23:52-2#

561 Interviewer: 多分ね、同時に聞き取りながら話すことは難しいと思います。 #01:23:58-0#

562 NHK1: それもね、それ結構訓練で、リスピーカーの人はそこは苦手だってという人はね、い
563 ないですですね。そう結構早くからみんながそれふつうにできてる人が多いです。僕なんか
564 も八か九年ぐらいいますけど、リスピークなんかできないですね、適正があるんじゃないかな
565 あという思いますけど。後リスピークがやっぱり言語能力が高くないとできないと思うのが、
566 文章をやっぱり頭で組み合わせていくことができないとはいけないと思うかな。話しているニ
567 ュアンスを伝えるために惑わしい文章まこういう風に言うのいいことでまあ、同作に近いかも
568 しれないけどもちゃんと文章にしたげられるかな。だから、あの、今字幕キャスターが多い

569 は、昔リポーターがやっていた人とか、アナウンサーをやっていたことで、リポーターができ
570 る力があるのですごくあの字幕キャスターにとっては重要な能力かもしれないですね。 #01:25:17-
571 3

572 NHK1: 放送前のリスピークのための準備って結構みなさんが台本しっかり読み込んでええと
573 分からないことちゃんと調べてくれます。で、ええとその周辺の言葉も結構拾ったりしてま
574 すね。テーマが何だろう例えばコロナウイルスだったらコロナウイルスの最新の言葉をバーと
575 メモして、あの、リスピークしてみて、出ない言葉があったら校正者にこの言葉が出ないから
576 ちょっと辞書に入れてというのは、そういうことはしますね。それではけっこう疲れていま
577 すよね。ダイレクターがいるしなきゃいけないですけど、そう。日本におけるリスピークの未
578 来。なんかみんな怖がっていますよねやっぱりダイレクト認識。音声認識がどんどん進むとダ
579 イレクト認識になって自分的のしごとがなくなっちゃいんじゃないかというふうに思ってる人達も
580 ありますね。たぶん一番最初になくなっているのはリスピーカーだったと思いますけどね。あ
581 の書きちゃだめですよ。やっぱりダイレクト認識になってやっぱりリスピーカーになくなっち
582 ゃうかなとおもいますね。それで、ダイレクトになったものこう力で修正いくっていうこの校
583 正者のほうが重要になってくるんじゃないかな。ただ日本語は難しいので中々ダイレクトとい
584 うのは非常にまだまだ時間かかるんじゃないかな。先いうた通りニュースは一人のアナウンサ
585 ーがゆっくりしゃべるからダイレクトにできるんですが、複数の話者が一斉に話始めるとダイ
586 レクト認識は非常に難しいですね。クロストークとかそれは認識しないし、だれかがもの起こ
587 したらことも[unverständlich]紛らわしい。一応日本語かかとも認識が??でも今の技術はそ
588 ういたら音楽とかそういうのも一ケース技術があって音声認識がすごくよくなってるんです
589 けども、中々ダイレクトは。 #01:27:55-7#

590 Interviewer: でも、オーストリアではそのダイレクトがぜんぜんありません。 #01:27:59-4#

591 NHK1: ないのか? #01:28:00-0#

592 Interviewer: ないです。あの句読点がとても大切ですから、それもリスピーカーがた例えば、
593 文章が終わったらまるといいます。ですので、ダイレクトにできません。 #01:28:21-3#

594 NHK1: ああ、そうなんだね。それはいいですね。なるほど。 #01:28:29-3#

595 Interviewer: ですから、そういう不安がありません。 #01:28:29-0#

596 NHK1: 英語はほぼダイレクトが進んでるらしいですね。アメリカのほうかね、ダイレクトの
597 音声認識が。 #01:28:45-2#

598 Interviewer: 多分ネットとかがありますが。わかりませんけど。 #01:28:52-6#

599 NHK1: テレビがあれですよ。??があったなんかで入力してますね。アメリカは回線でつ
600 ないてだいたくで放送入力してるっていうのは、日本はあの放送局で必ず人が集まってそれで
601 出してますね。たくさん話しましたが、他はありますか?

602 Interviewer: 大体大丈夫だと思います。

Interview Faith 1

- 1 Faith1: 私はメールに書いてましたけど、フェイスっていう会社の、もです。そして普段は新
2 橋ってわかります？ #00:00:51-4#
- 3 Interviewer: ああ、わかります。 #00:00:54-0#
- 4 Faith1: 新橋に本社があるんですけど、そこにいるんです。ただ今は大阪にいます。 #00:01:02-
5 7#
- 6 Interviewer: そうですね。 #00:01:03-1#
- 7 Person: はい。大阪に新しく字幕センターを今作ってるんですね。 #00:01:09-4#
- 8 Interviewer: ああそうですね。 #00:01:10-4#
- 9 Faith1: うん、そこで今システムをいろいろ組み立ってて、ええ、ですので、今大阪にずっと
10 出張してます。 #00:01:19-8#
- 11 Interviewer: では、多分貴社であるリスピーク方式についてちょっと説明してください。 #00:03:06-
12 5#
- 13 Faith1: われわれのリスピーク方式の字幕っていうのは、メールに書いてたかな。2014年
14 に始めて放送したんです。で、 #00:03:21-0#
- 15 Interviewer: それはNHKとどういう関係がありますか。 #00:03:25-7#
- 16 Faith1: まったく関係ないです。 #00:03:28-0#
- 17 Interviewer: ないです。 #00:03:30-0#
- 18 Faith1: その時はスポーツチャンネル衛星放送のチャンネルでした。大体2012年ぐらいです
19 かね、ただいろいろ研究を初めて。われわれは独自で作りました。ソフトと。それで正しい音
20 声認識は中々作るのは難しいんで、音声認識のエンジンは買ってきます。買ってきて、その
21 エンジンを使って、音声認識のソフトウェアを使って、修正して、字幕を出すというところは
22 我々のエンジニアがソフトウェアを書きました。 #00:04:25-6#
- 23 Interviewer: それは amivoice というソフトになりますか。 #00:04:28-5#
- 24 Faith1: そうです。われわれが初めて放送に分野に amivoice を使ったんですね。始めて使いま
25 した。なぜ使ったかという、ご存じだと思いますけど、NHKさんはすごくこうエクスペン
26 シブな #00:04:53-1#
- 27 Interviewer: ああそうですね。 #00:04:51-8#

28 Faith1: すごい素晴らしい音声認識をつかっていらっしゃるんですね。ただ我々は小さい会社
29 ですから、そんな大きいお金は中々使えないんで、amivoice っていうのは比較的手頃な値段
30 が分かります？あの安い値段なんです。 #00:05:16-0#

31 Interviewer: うん、分かります。 #00:05:16-0#

32 Faith1: なので、その amivoice をつかってなんとか、リスピークを使ってリアルタイム字幕を
33 できないかという研究をしたんですね。 #00:05:24-2#

34 Interviewer: ああそうですか。 #00:05:27-7#

35 Faith1: はい、なぜかというと、やっぱり安く作れないとたくさんの人が使えないですね、で
36 すのでシステムも自分たちで作って、音声認識も安いものを買ってきて、何とかシステムも作
37 ったということになります。 #00:05:46-6#

38 Interviewer: じゃ、他の会社は今貴社のシステムも利用していますか。 #00:05:55-6#

39 Faith1: あの例えば、J スポツってご存じです？ #00:06:03-9#

40 Interviewer: 聞いたことがあります、はい。 #00:06:01-6#

41 Faith1: スポーツチャネルなんですけど、そこはわれわれのソフトウェアを放送局に持ってい
42 て、オペレーターも一緒についていて、そこで、字幕を作ってます。 #00:06:26-8#

43 Interviewer: でも、リスピーカーも他のの会社に勤めていますか。 #00:06:31-6#

44 Faith1: リスピーカーはそうです、ほかの会社がやっています。アナウンサーの卵。 #00:06:38-
45 4#

46 Interviewer: アナウンサー #00:06:41-6#

47 Faith1: 将来にアナウンサーになりたい人がリスピーカーをします。大体うちの機材を他で使
48 うというよりもわれわれは例えば大阪とか、千葉県とか、そういうところに字幕センターを作っ
49 て、そこで放送局と通信をしながら、字幕を送ってるという仕組みでやっています。わかりま
50 す？ #00:07:14-3#

51 Interviewer: それは、なんか、貴社のスタジオで放送してということですか。リーモ一ととい
52 う。 #00:07:27-4#

53 Faith1: リーモ一と。そうです。だから、NHK の仕事もリーモ一とでやってますよ。で、フ
54 ジテレビって知っています。 #00:07:37-0#

55 Interviewer: 知ってます。 #00:07:42-0#

56 Faith1: フジテレビもつないでて、通信でやってます、リーモ一とで。 #00:07:51-1#

57 Interviewer: フジテレビもリスピーカー方式を利用していますか。 #00:07:56-3#

58 Faith1: いいえ、フジテレビの社内、フジテレビの放送局の中では、別の業者さん別の字幕メ
59 ーカーがあれば、キーボードかな、キーボードを打つ字幕を作ってます。で、フジテレビの仕
60 事は例えば20パーセントとかをフェイスがやって。あとは80パーセントはフジテレビの局
61 内でやってます。 #00:08:35-2#

62 Interviewer: そうですね。じゃ、どういう番組はフェスがしますか。どういう種類の番組。 #00:08:39-
63 6#

64 Faith1: 種類はニュース、はいとか。いろいろあります。スポーツも中継とか、ゴルフとかバ
65 レーボールとか、サッカーとかそれとあの、フジテレビあの、歌番組 #00:09:04-9#

66 Interviewer: そうですね。 #00:09:07-5#

67 Faith1: 歌番組もやってますし。ほかにやってますかな。あとバラエティー番組とわかります。
68 #00:09:14-6#

69 Interviewer: うん。 #00:09:16-9#

70 Faith1: コメディアンさん出る。そうやります。 #00:09:21-3#

71 Interviewer: それは、ジテレビの話でしたが、NHKはどうですか、どういう番組ですか。 #00:09:26-
72 7#

73 Faith1: 今NHKは毎日夜のニュースをやってます。ニュースとあと、難しいな、解説放送って
74 いうて、解説放送っていうアナウンサーじゃない人が専門的なことを十分間話すんですね。
75 夜にやってるんですけど、そうです。今だったらコロナウイルスに詳しい先生が十分間話すと
76 か経済について話すとか。そういったものをやってますね。 #00:10:25-6#

77 Interviewer: それは全部あのリスピークですか。 #00:10:29-8#

78 Faith1: あのニュースはリスピークです。で、解説放送は原稿があるので、それをテキストに
79 して、テーク方式って分かりますか。テーク、あのしゃべたら出すというやり方ですけど、解
80 説するその、なんというかな、話す人が間違うんですね。原稿じゃないことを言うんです。そ
81 うしたらリスピークに切り替えます。 #00:11:08-0#

82 Interviewer: その時は修正者とかいますか。 #00:11:10-2#

83 Faith1: ああ、いません、一人です。一人か二人ですね。 #00:11:14-3#

84 Interviewer: 一人です。そうか。 #00:11:17-5#

85 Faith1: 一人か二人何だけど、フェイスではリスピークができる人がいつもバックアップして
86 ます。だから原稿通りしゃべらなかつたらもうリスピークで話しますね。もっとも切り替えが
87 難しいです。 #00:11:42-6#

88 Interviewer: そうですね、じゃ、リスピーカーが一人だけですか。ああそうですね。 #00:11:56-
89 8#

90 Faith1: ああ、あの基本的なちょっと話をしていますか。 #00:11:58-4#

91 Interviewer: はい、どうぞ。 #00:12:01-1#

92 Faith1: われわれは amivoice を使って、音声認識にリスピーカーが話してテキストを作ります。

93 で、そのテキストっていうのは、認識率がわかります？間違え #00:12:17-7#

94 Interviewer: 認識率、分かります。 #00:12:19-6#

95 Faith1: それが、NHK はほとんど 99 パーセント言い訳ですね。オッケですね。でも、amivoice

96 は 95 パーセントぐらいしか正解しないです。でも、われわれは音声認識って必ずミスが置か

97 すっていう風に考えてるんですね。だから、ミスをするんだったらそのミスを人間、人を直し

98 ましょうという考えでシステムを設計したんで、われわれのシステムは amivoice で音声認識

99 したやつを修正者が三人とか、四人とか #00:13:09-9#

100 Interviewer: 三人 #00:13:12-3#

101 Faith1: 三人、四人で修正します。 #00:13:19-8#

102 Interviewer: それはいつも三人か四人ですか。番組の種類に違いますか。 #00:13:24-3#

103 Faith1: そうです。その通りです。 #00:13:26-7#

104 Interviewer: あ、そうですか。 #00:13:27-5#

105 Faith1: はい。番組の種類によります。 #00:13:32-0#

106 Interviewer: ちょっと、説明することができますか、どの番組はどうですかという。 #00:13:35-
107 6#

108 Faith1: ええとですね、ニュースとかバラエティー番組は三人か四人でやります。ただスポー

109 ツ番組とか、テレビショッピングは修正は一人か二人です。 #00:14:03-8#

110 Interviewer: ああそうですか #00:14:07-4#

111 Faith1: 理由は、ニュースは我々は百パーセント出さなきゃいけないです。しゃべったことは。

112 で、バラエティーも、バラエティーは百パーセントは出さないですけど、コメディアンがしゃ

113 べりやつね、百パーセント出さないけど、しゃべる数はすごく多いんで、リスピークした後に

114 ミスも結構字幕が、数が多いんですね。なので、人数も大きいしてます。ただ、テレビショッ

115 ピングとか、スポーツとか、スポーツはリスピーカーが頭の中で、少しい要約します。要約分

116 かります？ #00:14:51-3#

117 Interviewer: ああ、うん。 #00:14:53-8#

118 Faith1: 簡単にします。なぜかというたとえばスポーツの場合にサッカーで右に蹴りましたと

119 いうった時に七秒後にも左に蹴ってる場合もある七秒後、10秒後字幕が遅れて出るので、そ

120 したら、見てるほうがちょっと混乱するので、そこはリスピーカーがそういうことは言わな

121 いとか、いうことをしてます。 #00:15:25-8#

122 Interviewer: はい。 #00:15:28-8#

123 Faith1: あと、テレビショッピングの場合は #00:15:33-2#

124 Interviewer: はい #00:15:34-7#

125 Faith1: テレビショッピングの場合は女性の二人とか、三人で、こう…たくさん話しますね。
126 #00:15:43-5#

127 Interviewer: うん #00:15:44-9#

128 Faith1: それで、ほとんどが…なんというのかな話ことば。 #00:15:51-6#

129 Interviewer: あ、そうですか。 #00:15:50-5#

130 Faith1: あの、そうだな。すごいバーとか、なんかすごいとか、きれいとか、そういうこと話し
131 ますね。だから、それは少しいまとめて、きれい五回言うったら、一回しかいわないとか。た
132 だ、このペンがなぜすごいかよく書けるかということはちゃんと話すんですね。そういう風に
133 リスピーカーがようやくして話すので、修正者があまりたくさんいない、一人とか二人でや
134 ってます。その違いですね。 #00:16:30-9#

135 Interviewer: あの、リスピーカーのためその放送中は休憩とかありますか。交代は変わるとい
136 うことはありますか。 #00:16:42-0#

137 Faith1: 交代はほとんどしません。徳に NHK は CM がないですね。 #00:16:55-0#

138 Interviewer: CM? #00:16:55-0#

139 Faith1: CM, commercial。 #00:16:53-2#

140 Interviewer: ああ、CM #00:16:57-8#

141 Faith1: CM がないので、交代できないですね。だから、三十分とか、一時間はあの #00:17:11-
142 6#

143 Interviewer: それはいつも同じリスピーカーですか。 #00:17:14-7#

144 Faith1: ううん、いろいろな人がやります。 #00:17:16-7#

145 Interviewer: ああ、そうか、リスピーカーが変わりますか。 #00:17:19-3#

146 Faith1: いや、同じ人が一時間やります。 #00:17:26-5#

147 Interviewer: 一時間。ああそうですか、長いですよ。 #00:17:28-0#

148 Faith1: 長い、長いですね。でも、変わろうと思えば変わることができます。あの修正者は、
149 修正者、交代ができます。ソフトウェアの画面の中になんというかな、テキストをストップす
150 るっていうボタンがあります。 #00:17:56-7#

151 Interviewer: ああ、そうですか。 #00:17:59-5#

152 Faith1: それを押すことで四人いたら一人だけ休むんですね。で、例えば、バスルームに行く
153 とか、行って戻て来たらまた参加する。 #00:18:15-9#

154 Interviewer: そうですか。 #00:18:17-1#

155 Faith1: そういうことができるようにソフトウェアを作ってます。 #00:18:24-3#

156 Interviewer: うん、はい。あの、では、多分新しい人員を探す時は、何が一番大切な能力と資格ですか。リスピーカーのために。 #00:18:40-7#

158 Faith1: はい。リスピーカーは特に資格は聞いてません。いりません。我々のリスピーカーは
159 今何人かな、リスピーカーは10人以上いますけど、、特に大学とか、専門学校とかどこにで
160 っても大丈夫です。正しい、大事なのは、活舌。わかります? #00:19:28-2#

161 Interviewer: わかります。 #00:19:28-2#

162 Faith1: それが、いい人をできるだけ選んでます。 #00:19:39-3#

163 Interviewer: あの、テストとかありますか。試験とかその会社入るのために。 #00:19:46-9#

164 Faith1: ああ、簡単なあの基礎的なテストをします。例えば、最近の一般常識見たいなテスト
165 はします。ただ、ほとんどが面接ですね。面接で活舌がいいとか、声が出せるかどうかいう
166 ので選んでます。男性、女性も関係ないです。 #00:20:26-1#

167 Interviewer: それはいいと思います。 #00:20:26-0#

168 Faith1: はい。 #00:20:30-6#

169 Interviewer: はい。じゃ、その聞きながら話すことはみんなだいたい最初からできますか。
170 #00:20:44-7#

171 Faith1: できない。 #00:20:47-4#

172 Interviewer: じゃ、教育はどうですか。 #00:20:46-1#

173 Faith1: 教育はほとんど練習だけです。その聞きながら話すっていうことは、例えば amivoice
174 の場合は自分の声をずっと教え込まなきゃいけないですよ。そしたら、レベルがどんどん上が
175 っていくんですね。データが上がっていきます。それをやることによって、聞きながら話すっ
176 ていうことをになれるようにしてまわってます。ただ、これ、できない人もいます。でき
177 る人がそうだな、70パーセントぐらいかな。 #00:21:44-1#

178 Interviewer: そうですか。確かに難しい仕事ですね。 #00:21:47-6#

179 Faith1: 難しいですね。でもね、若い人がいいですね。 #00:22:05-4#

180 Interviewer: そうと思いますか。 #00:22:07-5#

181 Faith1: そうですね。今まで、うちの会社はパートタイムの女性をたくさん仕事してもらって
182 いるんですよ。でも、やっぱり、40代、50代は中々難しいですね。私もできません。笑
183 #00:22:36-6#

184 Interviewer: じゃ、ちょっとだけ字幕の品質についてそして精度について説明していただけない
185 いでしょうか。例えば基準とかがありますか。 #00:22:49-0#

186 Faith1: 基準はないですね、日本は。 #00:22:53-0#

187 Interviewer: ないですか。 #00:22:53-7#

188 Faith1: はい。 #00:22:55-0#

189 Interviewer: じゃ、会社の中もないですか。 #00:22:57-7#

190 Faith1: 会社の中はいろいろテストをしています。例えばリスピークのテストをしてリスピーカ
191 ーになるには、なんだろう、リスピークをして認識が悪い人といい人がいます。たとえば、百
192 なら百の字幕作ってもらって、リスピークでその時の認識率があるところより上にならないと
193 互角じゃないです。そういう風にテストをやります。 #00:23:46-0#

194 Interviewer: それについてあの、計算式とかがありますか。 #00:23:51-8#

195 Faith1: 計算式はないです。 #00:23:56-4#

196 Interviewer: 精度率について、ないですか。 #00:23:57-8#

197 Faith1: あの、なんていうのかな。例えば、しっかりした、そうですね。実際のニュースを見な
198 がら聞きながらリスピークをさせます。そうしたら、ちゃんとした、正解ニュースの正解のテ
199 キストとリスピークした後のテキスト音声認識のあとのテキストを比較します。その時に単語
200 の間違えの数がどれくらいあるかっていうので、あるラインを決めて判断します。それに合格
201 したらデビューをさせます。 #00:24:49-2#

202 Interviewer: ああそうですか。それは #00:24:58-1#

203 Faith1: それをやらないと認識率が悪いままデビューすると、リスピークはできるけど、修正
204 する人はとても大変でも、パニックになるので、しっかりやってますね。 #00:25:15-8#

205 Interviewer: それは、デビューのあとにも時々確認しますか。 #00:25:23-0#

206 Faith1: その通りです。してます。敵的にしてます。でもね、ある程度も数年たった人はやら
207 ないですね。逆に他の会社と比較しています。 #00:25:43-0#

208 Interviewer: そうですか。 #00:25:46-0#

209 Faith1: はい。 #00:25:49-1#

210 Interviewer: それは、どうしますか。 #00:25:54-0#

211 Faith1: 同じように正解のテキストを作って他の会社の間違え率と自分の会社の間違え率同じ
212 ように作業してそれを比較します。 #00:26:05-3#

213 Interviewer: 日本では、たくさんありますか、そういう会社。 #00:26:10-0#

214 Faith1: いや、ないですね。ええと、ね、五社か六社かぐらいですね。ちょっとこれは正確じゃ
215 ないですけど #00:26:24-4#

216 Interviewer: そのぐらい。 #00:26:25-4#

217 Faith1: そのぐたいですね。それは、あれですよ。リスピークだけじゃないですよ。 #00:26:34-
218 1#

219 Interviewer: キーボードとかも #00:26:34-3#

220 Faith1: キーボードとことも含めて #00:26:40-0#

221 Interviewer: キーボードはリスピークと比べて、キーボード多いですか。リスピークより #00:26:45-
222 3#

223 Faith1: 多いです。今はまだ、80パーセントがキーボードじゃないですかね。 #00:26:57-8#

224 Interviewer: それはなぜだと思いますか。 #00:26:59-0#

225 Faith1: 我々が結構リスピークが新しくて、、みなさんの会社は、他の会社は昔委から議会
226 が分かります、国会とか、あのものを議事録が分かります。エビデンス。 #00:27:23-3#

227 Interviewer: ああうん。 #00:27:26-6#

228 Faith1: そういったものを作る会社の人が、テレビの字幕を作っていくところが多いんですね。
229 だから、キーボードで始めてるところが多いんです。それも、実キーボードじゃなくてなん
230 か、特集なキーボードとか。 #00:27:44-3#

231 Interviewer: 見たことある、スピードワプロという見たことある。とてもすごいです。 #00:27:49-
232 0#

233 Faith1: すごいですよね。僕たちもそれを見た時これは、できないと思いました。彼女、彼ら
234 たちがそれをあの訓練、プラクティス、するのは、やっぱり三年とか四年とかかかるわけで
235 す。でも、僕たちはそんな時間もない、お金もないので、修正する人もリスピーカーもすごい
236 短い時間でできるようなシステムを作ろうということで、このリスピーク方式に決まりまし
237 た。 #00:28:25-0#

238 Interviewer: その、リスピーカーのトレーニングは平均的にどのぐらいかかりますか。 #00:28:34-
239 6#

240 Faith1: リスピーカーはそうですね。三か月でできる人もいます。 #00:28:42-9#

241 Interviewer: はい。 #00:28:49-0#

242 Faith1: 修正は二か月ぐらいできちゃいますね。だからとても速いです。例えば、大学生がイ
243 ンタンで夏休みからたまにくるんです。で最後にその時に四月に入社するときにはデビューがで
244 きます。 #00:29:32-2#

245 Interviewer: そうですね。速いですね。 #00:29:35-1#

246 Faith1: そのぐらい速いですね。リスピーク方式は本当お金がかからないですね。練習するの。
247 #00:29:45-4#

248 Interviewer: うんはい。多分字幕の遅延についてちょっと話しましょうか。それについて規定
249 とかありますか。 #00:30:02-0#

250 Faith1: 日本はないですね。 #00:30:07-0#

251 Interviewer: でも、普通にどのぐらいですか。 #00:30:10-0#

252 Faith1: 遅延は10秒、10秒ぐらいですかね。8秒から10秒ぐらいだと思います。 #00:30:21-
253 5#

254 Interviewer: はい。そして、字幕の速さと正しいさのどちらはより大切だと思いますか。 #00:30:35-
255 4#

256 Faith1: 正確さです、正しいさです。はい。ぜったい正確さです。 #00:30:48-9#

257 Interviewer: そして元の言葉り通りの字幕を作ることはどのぐらい重視ですか。本当に同じこ
258 とになりますか。 #00:31:10-3#

259 Faith1: 同じ、まったく同じように。 #00:31:12-7#

260 Interviewer: うん。 #00:31:14-7#

261 Faith1: 特にNHKとか、ニュース関係はアナウンサーが言うてる通りに字幕を作ります。かな
262 り大事です。 #00:31:26-5#

263 Interviewer: でも、テキストを省略することもあるとおもいますか。それはどうですか。 #00:31:33-
264 2#

265 Faith1: ニュースはないです。 #00:31:34-9#

266

267 Faith1: ニュースはない。でもスポーツとかはありますかはい、スポーツはあります、バラエ
268 ティーもあります。 #00:31:50-7#

269 Interviewer: それはどうしてですか。バラエティーは話ことが多いの… #00:31:49-0#

270 Faith1: 話ことばとあとクロストーク話さる時があります。その時は今話してるストーリーで
271 大事な人を選択します。 #00:32:13-9#

272 Interviewer: はい。では、あの、多分フィードバックについて、その字幕の退所者からフィー
273 ドバックとかをもらうことができますか。 #00:32:27-7#

274 Faith1: どういう内容ですか。 #00:32:29-9#

275 Interviewer: 例えば聴覚障害者から、高齢者から、 #00:32:33-2#

276 Faith1: はい、はい。聴覚障害者から我々の会社が直接もらうことはないですけど、NHK 例え
277 ば、須藤さん、まつこまさん、知ってます？ #00:32:52-0#

278 Interviewer: いいえ。 #00:32:55-1#

279 Faith1: ああ、須藤さんしか知らないんですか。 #00:32:54-3#

280 Interviewer: はい。 #00:32:57-1#

281 Faith1: ああ、あそこのグローバルメディアサービスの人たちが、例えば聴覚障害者団体に質問
282 したり。あとはそういう会議も開いた時にわれわれも出席します。 #00:33:23-6#

283 Interviewer: ああ、面白いです。どういうフィードバックがありましたか。 #00:33:31-8#

284 Faith1: そうですね。 #00:33:34-8#

285 Interviewer: 最近は #00:33:36-9#

286 Faith1: 最近ね、私は行かなかったんだよ。でも、話を聞いたんだけど、面白いのはあの字幕
287 に「ええ」を入れるほしい、それとか、あと「いいね」、とかね。 #00:34:00-7#

288 Interviewer: ああそうですか。全部ですか。 #00:34:01-6#

289 Faith1: それとか、あとなんだったかな。性格な字幕を聞きたい。あと、日本はいろを付けっ
290 てるんですね。 #00:34:26-2#

291 Interviewer: うん。 #00:34:31-5#

292 Faith1: あまりいろはいらないんじゃないかとか。 #00:34:31-5#

293 #00:34:31-5#

294 Interviewer: そうですか。 #00:34:27-6#

295 Faith1: うん。 #00:34:33-9#

296 Interviewer: じゃ、名前だけ書くとか。 #00:34:34-2#

297 Faith1: 名前を書くのでいいんじゃないかとか。 #00:34:39-4#

298 Interviewer: そうですか。 #00:34:43-5#

299 Faith1: ま、でも、聴覚障害者っていろんな人がいますね。 #00:34:48-3#

300 Interviewer: そうですね。 #00:34:48-3#

301 Faith1: 例えば、生前失聴と分かります。生まれた時から聞こえない人。 #00:34:56-0#

302 Interviewer: ああ、うん。 #00:34:58-8#

303 Faith1: で、途中で聞こえなくなった人。 #00:35:00-0#

304 Interviewer: はい。 #00:35:01-4#

305 Faith1: とか、年を取って聞こえにくくなった人。そういう人はやっぱり字幕を見てると思う
306 んですけど、その人たちでいろんなちがった意見がでるみたいですね。 #00:35:19-8#

307 Interviewer: 字幕の速さ、あの例えば 「ああ早くて読めない」という、文句がありました。
308 #00:35:31-9#

309 Faith1: いや、あのそれはあまり聞きません。文句はないけど、 #00:35:37-3#

310 Interviewer: そういう意見とか。 #00:35:40-8#

311 Faith1: 意見は前、昔聞きましたけど、大体リアルタイム字幕、生放送字幕っていうのは NHK が
312 3 秒です。表示する時間が。でも、民放は 2 秒です。 #00:36:07-9#

313 Interviewer: 民放？ #00:36:07-9#

314 Faith1: 民放ってフジテレビとか。 #00:36:10-6#

315 Interviewer: ああそうですか。 #00:36:12-7#

316 Faith1: TBS とか。 #00:36:15-8#

317 Interviewer: そうですね。 #00:36:18-6#

318 Faith1: そこは 2 秒なんです。 #00:36:19-4#

319 Interviewer: あ、違いますか。 #00:36:21-6#

320 Faith1: はい。NHK はルールが厳しくて、アナウンサーがゆっくり話さないといけないですね。
321 でも、そのフジテレビとか commercial やってる民放放送はあまりルールがない、速くしゃべ
322 ちゃうんですよ。 #00:36:54-1#

323 Interviewer: そうですね。 #00:36:56-0#

324 Faith1: そうすると、字幕がたくさんになるんでゆっくり出しているとどンドンどンドン遅れち
325 ちゃうんですね。だから、今は 2 秒ですね。 #00:37:09-6#

326 Interviewer: うん。 #00:37:13-4#

327 Faith1: はい、それはテレビ局が決めます。日本語で、最大 3 5 字ぐらいですね、一つの字幕
328 が。 #00:37:28-3#

329 Interviewer: では、大体そういう悩み、スピードについての悩みがありませんか。 #00:37:31-
330 3#

331 Faith1: ないと思います。あまり聞かないですね。 #00:37:39-0#

332 Interviewer: はい、じゃ、多分みんな早く読めますね、日本で。 #00:37:42-6#

333 Faith1: そうですね。なんか、結構みんな聞こえない方って読むのが早いみたいですね。 #00:37:48-
334 8#

335 Interviewer: そうですね。オーストリアでは、手話を使用している方は時々読むことは難しく
336 なります。そういう人は問題があることあると思います。 #00:38:05-1#

337 Faith1: はい、手話を使っている人ですね。 #00:38:05-0#

338 Interviewer: はい。 #00:38:07-8#

339 Faith1: なるほど。 #00:38:13-1#

340 Interviewer: それは、日本でどうですか。わかりますか。 #00:38:12-2# #00:38:33-8#

341 Faith1: わかりません。 #00:38:18-4#

342 Interviewer: 多分そういう団体に聞くほうがいいですね。 #00:38:29-7#

343 Faith1: そうですね。 #00:38:28-8#

344 Interviewer: はい。字幕はいつも日本語から日本語にだけですか。ほかの言語から日本語に翻
345 訳した字幕がありますか。 #00:38:39-0#

346 Faith1: 今はありません。でも、将来的に例えば、日本語聞いて英語で出すっていうのは簡単
347 にできると思います。 #00:38:53-4#

348 Interviewer: 簡単。 #00:39:00-6#

349 Faith1: 簡単です。同時通訳があるじゃないですか。 #00:39:04-6#

350 Interviewer: そうですね、でも同時通訳も簡単じゃありません。 #00:39:08-8#

351 Faith1: ただ、同時通訳は簡単じゃないですけど、できる人はたくさんですね。 #00:39:15-5#

352 Interviewer: たしかに。 #00:39:13-7#

353 Faith1: 後は、英語を音声認識すればそれを字幕にすることはできるし、ただ修正する人もい
354 りますね。 #00:39:28-5#

355 Interviewer: そうですね。 #00:39:30-0#

356 Faith1: だから、今と同じシステムで、音声認識のソフトウェアを買えることで、日本語から
357 英語とか、日本語からオーストリア語とかというのたのになるかもしれません。 #00:39:43-1#

358 Interviewer: では、今そういう計画がありますか。そういうソフトを作る。 #00:39:55-7#

359 Faith1: 今はないですけど、もしかしたら、数年後に万国博覧会があります。分かります。 #00:40:09-
360 2#

361 Interviewer: 分かります。 #00:40:05-6#

362 Faith1: それが、あるので、それに向けてやるかもしれません、まだ決まってません。 #00:40:15-
363 3#

364 Interviewer: はい。では、生字幕の未来についてどう思いますか。 #00:40:25-8#

365 Faith1: 今ですね、amivoice をつかってますけど、これを例えばグーグルとかスマホはすごく優
366 秀ですね。これでできますね。でも、これはいつも通信でつながってるからできるんですね。
367 だから、放送に使うのは僕たち心配なんです。だから amivoice はなぜいいかと、システムの
368 中に入っているって音声認識のソフト。だから、切れることもないですね。 #00:41:24-1#

369 Interviewer: はい。 #00:41:25-9#

370 Faith1: だから、ま、将来的にそのこういった、スマホで使ってるような音声認識につながぐ
371 こと許されれば、もっと優秀な音声認識を使っていけると思います。四人いる修正者を一人に
372 したいと考えます。 #00:41:49-6#

373 Interviewer: そうですか。オーストリアでは Dragon naturally speaking というソフト使用されて
374 います。聞いたことがありますか。ドラゴンスピーチ。 #00:41:59-5#

375 Faith1: ドラゴンスピーチ。あ、聞いたことがあります。それは、オーストリア #00:42:09-5#

376 Interviewer: テレビで。 #00:42:12-7#

377 Faith1: オーストリアの製品？ #00:42:14-2#

378 Interviewer: ううん、そうじゃないです。オーストリアで使用されているプログラムです。オ
379 ーストリアの公共テレビで。 #00:42:21-7#

380 Faith1: あ、そうですか。ドラゴンスピーチ。これもそんな高くないですね。 #00:42:28-7#

381 Interviewer: はい。分かりませんがそうだと思います。 #00:42:38-2#

382 Faith1: それ公共放送ですか。あれです。 #00:42:40-3#

383 Interviewer: はい。ヨーロッパの中によく使用しているものだと思います。 #00:42:52-4#

384 Faith1: そうですか。それもシステムの中に入れるやつですね。 #00:42:55-7#

385 Interviewer: はい。 #00:42:57-3#

386 Faith1: 通信は使わないやつですね。 #00:43:00-1#

387 Interviewer: はい。 #00:42:58-1#

388 Faith1: なるほど。 #00:43:00-5#

389 Interviewer: 日本語もできますと思います。ドラゴンスピーチ。 #00:43:01-3#

390 Faith1: はい。ありました。でも、前試したけど、あのいくつか試したんです。その中で一番良
391 かったのは amivoice なんですね。 #00:43:18-9#

392 Interviewer: うん。そうですね。 #00:43:24-5#

393 Faith1: 将来的にはもっともっと AI 人工知能が進んで言って、認識率がもほぼほぼ百パーセン
394 トになる可能性があると思うんですけど。でも公共放送とかは、間違えたらいけないとおも
395 うんですね。ミスがあったら。だから音声認識でとんでもない間違えををすると思いません。人
396 間は間違う時にちょっと間違うですね。でも音声認識でなんか、何だろう間違えますよね。と
397 んでもない間違えますね。だから、そういうのはやっぱり注意しなきゃいけないんで修正する
398 人は一人が必要かなと思ってます。 #00:44:26-0#

Interview Faith 2

1 Interviewer: こんにちは。 #00:00:16-4#

2 Faith2: こんにちは #00:00:30-3#

3 Interviewer: 聞こえますか。 #00:00:32-7#

4 Faith2: 聞こえます。 #00:00:34-4#

5 Interviewer: はい、私もうそうです。よろしく願います。 #00:00:38-5#

6 Faith2: 願います。 #00:00:40-4#

7 Interviewer: では、今日を忙しいところありがとうございます。 #00:00:48-1#

8 Faith2: いいえ、いいえとんでもないです。 #00:00:50-2#

9 Interviewer: いくつかの質問がありますが、でも最初に増田さんは、増田さんはもうリスピー
10 カーと務めていますか。今は。 #00:01:04-4#

11 Faith2: はいそうです。 #00:01:04-2#

12 Interviewer: いつぐらい働き始めましたか。リスピーカーと。 #00:01:09-2#

13 Faith2: 2012年なので、今8年間やってます。 #00:01:18-1#

14 Interviewer: ああ、そうですか。長いですよ。あの、ちょっと個人的な情報なんですが、多分
15 職歴と学歴について説明いただけないでしょうか。 #00:01:37-4#

16 Faith2: はい。専門学校って分かりますか。 #00:01:43-6#

17 Interviewer: 分かります。 #00:01:44-0#

18 Faith2: あの、ま、短い大学みたいなものに通っていました。そこで、テレビとかのそういう
19 業界のことを勉強しました。で、あとに今自分がいるフェイスという会社に勤めだしました。
20 #00:02:08-3#

21 Interviewer: そうです。でも。どうしてリスピーカーになりたがっていましたか。 #00:02:18-7#

22 Faith2: 最初はその会社の方針です。上からこう命じられて、ちょっとこれでチャレンジして
23 みてっといわれて。このリスピークということをやっていました。 #00:02:40-6#

24 Interviewer: はい、あのそのリスピークの仕事のための教育はそれはフェイス社で行われまし
25 たか。 #00:02:59-5#

26 Faith2: 今きこえてますか。 #00:03:12-6#

27 Interviewer: あれですよ。練習方法ですよ。 #00:03:13-8#

28 Faith2: はい。 #00:03:17-6#

29 Interviewer: でも、それも今の会社に勤めだしたからです。で、何でしょうね。いろいろなテ
30 レビ番組を見ながら、ずっと音を聞きながら、自分で声に出すという方法で独学で自分で学び
31 ました。 #00:03:40-2#

32 Faith2: そうですか。それは練習、本当のリスピークの練習だけでした、他の教育と練習があ
33 りましたか。 #00:03:50-4#

34 Interviewer: でも、いろいろな番組をこう見たので、例えば政治の問題だったり、バラエティ
35 番組だったり。後はなんだろうな。あとスポーツだったり。いろんなジャンルのいろいろな単
36 語を勉強をする必要がありました。 #00:04:19-3#

37 Faith2: たしかにそうです。大体そういう教育についてどう思いますか。なんか、もっと詳し
38 い説明をもらったらいいというところがありますか。 #00:04:35-4#

39 Interviewer: 今自分はそのいろいろなテレビ局と仕事をしています。で、そこからでもいろん
40 な情報をもらえます。なので、そういう情報があれば、自分でリスピークして対応することが
41 できています。なんですけど、やっぱり基本的な知識がないと、対応するのが難しいですね。
42 #00:05:20-9#

43 Faith2: はい。 #00:05:22-5#

44 Interviewer: なんで、やっぱり自分で、自らこう学びにいかないとちょっと難しいかな、と思
45 いますね。いろんなことを率先して吸収して難しいかなと思います。 #00:05:41-2#

46 Faith2: それは教育の中にありませんでした、そういう情報? 多分時間がないですが #00:05:41-
47 2#

48 Interviewer: そうですね。今この会社に勤めだして、リスピークってやり始めたんのは自分が
49 初めての人間になったね、一人目の人間何です。 #00:06:03-9#

50 Faith2: はい #00:06:05-1#

51 Interviewer: なので、自分よりあとに会社に入った人にはそういうふうな教育をしています。
52 #00:06:11-5#

53 Faith2: そうですか。 #00:06:12-1#

54 Interviewer: 自分から学びに行くんだよってことを教えてあげています。 #00:06:19-2#

55 Faith2: はい。今はどういう種類の番組を字幕付与していますか。 #00:06:32-3#

56 Interviewer: さきほども言いましたが、ニュース番組、スポーツで。スポーツっていうと野球
57 とか、フットボール、バスボールとか。あとはなんですか。でもそうですね、スポーツはその
58 ぐらいですね。後はお笑い番組、バラエティ番組ですね。 #00:07:08-5#

59 Faith2: はい #00:07:10-5#

60 Interviewer: 増田さんはそれが全部しています。専門とかはありませんか。 #00:07:17-3#

61 Faith2: 専門はなくて、全部の番組に対応できるようにしています。 #00:07:24-2#

62 Interviewer: それはリスピーカーさんの皆さんそうですか。全部をやっていますか。 #00:07:28-
63 7#

64 Faith2: やっぱ、ちょっと向き不向きがあります。例えばどうしても展開が早い番組だと、
65 ちょっと入ったばかりの社員にはまだやらせてないとかはありますね。ただ、最初的には全員
66 が全部対応できるように教育はしています。 #00:07:51-5#

67 Interviewer: はい。多分なんか、バラエティーとかは多分難しいと思いますね。 #00:07:56-2#

68 Faith2: そうですね。 #00:07:59-1#

69 Interviewer: 他に、どういう種類が一番難しいと思いますか。 #00:08:05-9#

70 Faith2: 政治的な話ですね。 #00:08:08-7#

71 Interviewer: それはどうしてですか。 #00:08:11-7#

72 Faith2: 結構、センシティブというか。いろいろな違う思った人を見てるものなんで。誤った
73 情報を出してしまうとやはりクレームの対象になっちゃうんですね。 #00:08:25-2#

74 Interviewer: そうですか。あの、クレームはよくありますか。 #00:08:30-4#

75 Faith2: いや、放送局には来てるみたいなんですけど、一応私たちが務めている会社には、あ
76 まりこういわれることはないです。 #00:08:41-6#

77 Interviewer: そうですか。では、リスピーク方式だけで字幕を付与していますね。それは毎日
78 やっていますか。 #00:08:54-3#

79 Faith2: でも、そうですね。ほとんど毎日やっています。 #00:09:07-7#

80 Interviewer: その字幕以外にほかの仕事はありますか。 #00:09:13-2#

81 Faith2: 今私がやっているのは収録されたすでに出来上がっているものにも字幕を作っています。
82 #00:09:34-6#

83 Interviewer: 事前放送ということ #00:09:34-9#

84 Faith2: そうですね、事前放送、事前収録前に、ろくがしておいて例えばドラマとかリアルタ
85 イムではない、生放送ではない、そういうのにも一応字幕を作っています。 #00:09:55-9#

86 Interviewer: それは他のリスピーカーもこういう仕事をしていますか。 #00:09:59-1#

87 Faith2: 今、そうですね。実はここは事務所を作ったばかりです。で、新しく社員を雇ったば
88 かりですね。 #00:10:19-0#

89 Interviewer: ああ、それは大阪の #00:10:22-0#

90 Faith2: そうです。 #00:10:22-4#

91 Interviewer: そうですか。うん。 #00:10:25-6#

92 Faith2: 大阪なので、他のはいつてきたばかりの社員は主にトレーニング今が結構マインの
93 仕事になっています。 #00:10:39-2#

94 Interviewer: そうですか。分かりました。あの、大体リスピーカーの就業時間、働いている時
95 間はどのぐらいですか。パートですか。フルタイムですか。 #00:10:50-6#

96 Faith2: フルタイムですけど、なんだろうな。実際こうしゃべる時間だけで言うと、一応三時
97 間ぐらいですがね、 #00:11:12-6#

98 Interviewer: じゃ他に他の時準備とかをしていますか。 #00:11:17-4#

99 Faith2: そうですね、放送前にはどうしても準備が必要です。 #00:11:22-4#

100 Interviewer: それはどういう準備ですか。 #00:11:23-3#

101 Faith2: その日の放送内用を確認したり、後はその日一日で起きた事件とか、事項を何でしょ
102 う。。。人の名前を出す必要があるですね。被害者だったり、容疑者だったりなので、そうい
103 う人の同士も日本語漢字がたくさんあるので、そういう漢字の調べみたいなんです。 #00:11:56-
104 3#

105 Interviewer: 分かりました。その字幕システム、あの音声認識のシステムのトレーニングもし
106 ますか。例えば音声認識がよくわかるのために。そういうことがありますか。 #00:12:16-1#

107 Faith2: 今のさっきいた新しくはいった社員が今トレーニング中です。で、それはなんでかと
108 いうのももちろん自分が聞いてしゃべるというのをうまくなるため何ですけど、実は音声認識の
109 アプリがあるですけど、それにやっぱり声覚えさせないとはいけないでsね。で、一応その今
110 トレーニングを兼ねています。 #00:12:50-2#

111 Interviewer: はい、多分あの、例えばこのシステムの語彙にはいてない単語とかもありますね。
112 それもそれについても練習がありますか。 #00:13:05-6#

113 Faith2: それはそうですね。放送のさき言った前の時間の準備の時間に、例えば容疑者の名前
114 とか言うってみて、当選でないですね。なので、その時に我々は辞書っていうんですけど、デ
115 イクショナリ辞書みたいなもの作り育てそれをその放送前の時間に作るという作業がありま
116 す。 #00:13:32-4#

117 Interviewer: 増田さんはリスピーカーだけですね。校正者もやっていますか。 #00:13:41-4#

118 Faith2: 自分もやってます。 #00:13:41-5#

119 Interviewer: よくやってますか。 #00:13:48-2#

120 Faith2: そうですね。一緒にやっています。 #00:13:53-3#

121 Interviewer: そうですか。いつも交代しますか。リスピーカーと修正者。 #00:13:57-2#

122 Faith2: そうですね。リスピーカーと修正者は基本的にはその日一日で交代することはないで
123 す。一日どっちかの作業に専念するという感じです。 #00:14:15-0#

124 Interviewer: でも、他の日は交代しますか。 #00:14:18-1#

125 Faith2: そうそうそう。 #00:14:20-9#

126 Interviewer: そうですか。ああ #00:14:20-9#

127 Faith2: 例えばさっき言ってたのほぼ毎日やっているの、例えばこの増田だけにリスピーク
128 してしまうと例えば自分が何が体調不良とか事故とかそれこそ起きたことに交代にだれでも
129 かわれるようにしなきゃいけないですね。あまりよくない時でも、なので、その本当はその日
130 事にリスピーカーを作るという漢字ですね。 #00:14:50-8#

131 Interviewer: では、修正者のためにどういうトレーニング、どういう教育がありますか。 #00:14:58-
132 7#

133 Faith2: 修正者はそれこそはキーボードを使っています。なので、タイピングが早くないとい
134 けません。なので、そういうタイピング日々うってんもらってん練習したりとか後はさっきも
135 言ってたんですけど、いろんな人の名前を日本語が漢字があるので、ただし漢字が出てくるよ
136 うにそれちょっと変換をするという作業が発生しています。それはトレーニングをしています。
137 #00:15:46-3#

138 Interviewer: 会社に入る前にテストとかがありますか。 #00:15:56-4#

139 Faith2: そうですね。ただ、それを適正を見極めるみたいなテストはないです。一応我々の方
140 式だと比較的だれでもできるものです。トレーニングしてもらえば。なので、例えば特別、こ
141 の人はしゃべるのが上手とか、さっき言ったタイピングが早いとかがオッケではなくて、入っ
142 た後に全然トレーニングしようようっていう感じですね。 #00:16:37-6#

143 Interviewer: 普通の面接だけですか。 #00:16:44-7#

144 Faith2: そうですね。後はその何でしょうね。一般的な次式があるかどうかを一応その面接の
145 会話で探ったり、後は一応スポーツが好きだったりニュースが、テレビを見るとか好きかとか
146 という質問はしますね。で、さっきもいったけど、自分で情報へに行くのが好きかどうかとい
147 うのはちょっとみたいはしています。 #00:17:20-8#

148 Interviewer: わかりました。そうですね。リスピーク中にどういう問題はよくありますか。
149 #00:17:35-7#

150 Faith2: リスピーク中 #00:17:36-9#

151 Interviewer: 問題と間違えとか #00:17:39-8#

152 Faith2: そうですね。 #00:17:45-2#

153 Interviewer: 何が一番難しい状況とか #00:17:44-4#

154 Faith2: 日本だとたまに大きい地震が起きたりします。 #00:17:54-0#

155 Interviewer: そうですね。 #00:17:56-5#

156 Faith2: 災害というですけど、地震だったり、台風とか、その時というのは番組が急に生放送
157 になったりするんですね。で、例えばすごい長い時間生放送急にテレビ局のほうからやってく
158 らいか依頼が来ることがあります。そうすると番組中にコマーシャルにいけないんですね。それ
159 は休憩の時間がなくなって、コマーシャルが入らないと。長時間リスピークしないと行くなくな
160 っちゃうんですね。それがさっき言ってた三時間とかじゃなくてもっと倍の時間とか。 #00:19:01-
161 8#

162 Interviewer: そうですね。 #00:19:02-6#

163 Faith2: 飲み物も飲まなかったり。あるいは、おなかがすいちゃったり、トイレに行けなかつ
164 たり。そういうやっぱり急なトラブルみたいなものにそういう放送しなきゃいけないのでわ
165 れわれね。そういう時が一番大変かもしれないですね。 #00:19:29-7#

166 Interviewer: はい。その時はリスピーカーを交代しませんか。 #00:19:33-5#

167 Faith2: さすがにさっき言ってた六時間とかものすごい長い時間になっちゃった時は交換しま
168 す。 #00:19:46-2#

169 Interviewer: どのぐらいの頻度に交換しますか。 #00:19:52-3#

170 Interviewer: 決まった頻度がありますか。 #00:19:58-3#

171 Faith2: これもでもやっぱり人によります。さっき言った急に体調が悪くなったということ
172 もあるので。急にトイレに行きたいとかどうしものどがかわいだとかという時はさっき言った
173 コマーシャルがはいったらすぐに別の人に交換する見たいな感じをすることもあります。 #00:20:22-
174 7#

175 Interviewer: コマーシャルがない時は #00:20:24-6#

176 Faith2: そうそうそう。番組のスキーマを縫ってやることはあります。 #00:20:37-4#

177 Interviewer: 大変みたいですね。問題や間違えを避けるためにどの戦略とかがありますか。

178 #00:20:50-0#

179 Faith2: さっきも言ってた放送の前の準備あそこはやっぱり念入りやるしかないですね。でき
180 るだけ正しい情報を出してあげる。なので、そのために変換のわれわれは中に入ってるんですけ
181 ど、これはでも日本の多分キーボードにしかないんですけど正しい漢字が工務がいっぱいでっ
182 て来るんですね。なので、例えば自分の増田っていう漢字を出したり時はその漢字を使い続け
183 ていると漢字自体か上に上がってくるんですよ。 #00:21:54-9#

184 Interviewer: そうですか。よくあることは上に来ます。 #00:21:59-2#

185 Faith2: そうです。そういうことをやっていますね。 #00:22:05-8#

186 Interviewer: 多分きゅうに知らない名前がでて来ましてどうしますか。 #00:22:16-3#

187 Faith2: われわれの会社ではカタカナで出しています。 #00:22:27-3#

188 Interviewer: そうですか。私はこういう仕事ではストレスがたくさんあると思います。それは
189 どのように対処しますか。 #00:22:39-5#

190 Faith2: ストレスの対処はどうだろうな。難しい質問ですね。 #00:23:03-5#

191 Interviewer: そうですね。わかります。たくさんあるかと思いませんか。 #00:23:12-5#

192 Faith2: 生放送なので、やっぱりそうですね。ものすごいストレスをかかっているとおもいま
193 すね。ただ、どう対処するかというのは難しいですね。何て言ったらというかな。でも、やっ
194 ぱり本当に丁寧に準備しておくてことしないかな。 #00:23:44-1#

195 Interviewer: わかります。多分考えられないも大丈夫ですけど、難しい状況がありました時に
196 おぼえられますか。そういう経験がありましたか。 #00:24:07-8#

197 Faith2: なんだろう。でも、答えとして後いるかわからないですけど。パソコンのトラブルと
198 か #00:24:50-8#

199 Interviewer: そうですね。技術的な問題。 #00:24:54-3#

200 Faith2: そうです。それはわれわれだと解決できない問題が起きてしまうことが一番をなりま
201 したね。 #00:25:09-7#

202 Interviewer: 分かりますね。大変ですね。そうですね。多分他になんか言いたいことがありま
203 す、いつでもいいですが、多分ほかのテーマに行きましょうね。それは、私は句読点に興味
204 があります。どのように字幕で行われていますか。くとうてん。 #00:25:48-7#

205 Faith2: 句読点。字幕って制限があるですね。何文字っていうのは決まっています。 #00:26:22-
206 4#

207 Interviewer: 分かります。はい。 #00:26:23-6#

208 Faith2: なので、できるだけでもその制限ギリギリまで文字を使いたいんですよ。なんですけ
209 ど、われわれが使ってるひらがなってどうしてもつづちゃうと見つらくならんですね。でも、
210 いろんな単語に創造できてしまうんです。さっき句読点というのでいたくいてあげてます。見
211 ってる人にとって見やすいでだろうなってところに付けてあげてます。 #00:27:03-4#

212 Interviewer: そうですね。それはだれがいますか。校正者ですか。 #00:27:11-4#

213 Faith2: リスピークするといちおう自動的に生成されてます。 #00:27:23-7#

214 Interviewer: 自動的ですか。 #00:27:25-8#

215 Faith2: そうです自動的です。で、ただ、最終的な判断をするのは修正者です。 #00:27:33-0#

216 Interviewer: そうですね。それは話者の変わるはどのように表示しますか。色ですか。 #00:27:51-
217 9#

218 Faith2: 自分たちの会社ではオフマークというマークがあります。ええとね。今でチャットを
219 見れますか。 #00:28:12-4#

220 Interviewer: はい。 #00:28:15-0#

221 Faith2: このマーク何ですけど。 #00:28:21-1#

222 Interviewer: 分かります。はい。 #00:28:24-8#

223 Faith2: このマークが入ったら話者が変わったという。 #00:28:32-5#

224 Interviewer: 色はありませんか。 #00:28:35-5#

225 Faith2: われわれはまだ色はやれてません。 #00:28:40-6#

226 Interviewer: そうですね。それはもしスピーカーか修正者かだれの仕事になりますか。 #00:28:49-
227 6#

228 Faith2: いろですか。こういうシンボル #00:28:55-7#

229 Interviewer: マーク。修正者ですね。 #00:29:05-8#

230 Faith2: そうですね。修正者が直しながら入れて行きます。 #00:29:17-0#

231 Interviewer 分かりました。あとはね、多分ちょっとおかしい質問ですが、増田さんの意見によ
232 ると字幕の何割ぐらいは本当に言葉通りですか。確かに内容は同じですけど #00:29:43-8#

233 Interviewer: どうだろうな。90パーセント95パーセントぐらいはその放送に対しては出し
234 たいなっていう気持ちはあります。でも、どうしてもこのリスピーカーってリスピーク方式だ
235 と何人もが同時にしゃべる時があるですね。 #00:30:13-2#

236 Faith2: はい #00:30:14-6#

237 Interviewer: そういう時例えばやっぱり出せないで、片方の人の言葉しか出せないですよ。
238 なので、例えばこう同時に話してしまった時はどちらかの人は何も出なかったとなっちゃうん
239 ですね。だからまあそうですね。90パーセントくらいは出せてるかなあとと思います。 #00:30:40-
240 7#

241 Faith2: はい。話者は言葉を繰り返した時。例えば同じことは多分二つ目言いました時それは
242 どうします。略することしますか。 #00:30:59-8#

243 Interviewer:略すこともありますね。 #00:31:02-2#

244 Interviewer: 例えば。 #00:31:06-6#

245 Faith2: なんかこうなんだろうな。でも、ぜんぜん略すと思います。 #00:31:15-8#

246 Interviewer: そういう略すかどうかそういう はんだんすること #00:31:28-1#

247 Interviewer: はどうしますか。 #00:31:32-8#

248 Faith2: これもどうでしょうみて全部のつながりを見るしかないですね。日本語のようは文章
249 見ててなんで言ってるかなあ、判断する...難しいですね。 #00:32:02-2#

250 Interviewer: そうですね。難しい質問です、ごめん。 #00:32:05-2#

251 Faith2: いやいや...でも、割と二回言うっちゃう人っていうのは何かこう考えながら話してる人
252 が多い気がします。なので、同じこと二回言うっちゃうですなので。だから、リスピークの判
253 断で消してます。そういう時はいったん話してみてもたまたま同じこと言うたこの人思ったらしゃべ
254 らなかったりします。 #00:32:43-2#

255 Interviewer: そうですね。 #00:32:47-2#

256 Faith2: で、続きから話してます。 #00:32:49-9#

257 Interviewer: そうですね。リスピークの時に文章が終わりまで聞きましたか。その後にリスピ
258 ークをしていますか。 #00:33:01-3#

259 Faith2: いや、話す途中です。 #00:33:02-8#

260 Interviewer: 途中ですね。それは難しいですか。多分その文章の終わりがまだわからないから
261 すぐに代わる文章もありますので、それについてどう思いますか。 #00:33:22-0#

262 Faith2: どうな。ええと、難しいですけど、どういったらいいかな、基本的には話してる人が
263 話し出したら本当にすぐに一分を話し出してます。ただ、言葉の意味とか文章の内容はあまり
264 頭に入れないようにしています。音でしかとらえてないというか。意味を考えないようにして
265 います。 #00:34:13-3#

266 Interviewer: そうですね。 #00:34:15-7#

267 Faith2: そうじゃないと。さっき見たいに会話を聞いちゃうとなんか上手くしゃべるなかなる
268 んですね。あとまあでも音いうった通りに自分が再生するんだという感じでしゃべっていま
269 す。 #00:34:33-0#

270 Interviewer: そうですね。面白いです本当に。私はね、ちょっと同時通訳の経験がありますの
271 で、ちょっと多分こういうことが分かりますが違うですから、面白いです。リスピーカーにと
272 って一番大切な資格はどれとおもいますか。 #00:35:02-1#

273 Faith2: でも、勝節かもしれないです。 #00:35:07-1#

274 Interviewer: そうですね。 #00:35:09-1#

275 Faith2: 言葉をはっきりと話す能力 必要だとおもいます。もう一個いうとしたらさっきも言う
276 ったんですけど、幅広いいろんなことを知るかどうかとも求められると思います。 #00:35:28-7#

277 Interviewer: ほかの大切なことがありますか。 #00:35:31-7#

278 Faith2: そうだなあ。あとはでもやっぱりリスピークだとちょっと言葉を文章をどうしても作
279 るんないとはいけないあるんですね。頭で若干と考える時はあります。なのでやっぱりこ
280 の文章力というか、あとはその日本語を知ってるかどうかとも試されてしまいます。そういう
281 言葉があるかどうかの知ってないとやっぱりしゃべることができないですね。なので、そ
282 んな変ですね。 #00:36:23-4#

283 Interviewer: じゃ修正者はどうですか。 #00:36:37-7#

284 Faith2: 修正者はでもやっぱりさっきも言ったけど、タイピングそこは求められちゃうんです
285 ね。 #00:36:53-7#

286 Interviewer: 何が一番難しいと思いますかリスピークと修正どっちですか。 #00:37:06-8#

287 Faith2: 修正よりリスピークですね。リスピークです。 #00:37:14-7#

288 Interviewer: リスピークですか。どうして。 #00:37:14-8#

289 Faith2: リスピークのほうがすごい三時間しゃべり続けるんですよ。 #00:37:22-8#

290 Interviewer: そうですね。 #00:37:24-1#

291 Faith2: ものすごく疲れます。 #00:37:26-8#

292 Interviewer: はいたしかにそうです。三時間は本当に長いと思います。私同時通訳の時は多分
293 なんか 2 0 分 3 0 分だけです。本当にすごいです。 #00:37:41-2#

294 Faith2: いいえ、いいえ。 #00:37:48-3#

295 Interviewer: はいそれはも私のことなんですけど、リスピークと同時通訳にとって費用な能力
296 や資格が似ていると思いますか。ということは聞きましたか。最初は。 #00:38:14-0#

297 Faith2: 自分からそれと同時通訳のほうが難しいと思います。 #00:38:27-0#

298 Interviewer: 分からにけど、でも、オーストリアとヨーロッパの中にそういうことはよく言われ
299 ていますね。同時通訳とリスピークはにてるので、そういう質問があります。それは日本
300 でしょうか。 #00:38:38-6#

301 Faith2: 日本でも、でもそういうのは求められています。英語で放送してるものも日本語の字
302 幕にできたらっていうのは求められています。 #00:38:55-3#

303 Interviewer: それはありますか。 #00:38:58-3#

304 Faith2: あります。あります。 #00:39:02-3#

305 Interviewer: そうですか。でも、貴社でありますか。 #00:39:02-9#

306 Faith2: うちではまだできてないです。自分がまだ全然英語がわからないので、やっぱりそ
307 こは分かる人はいれば挑戦はするとおもいます。 #00:39:18-8#

308 Interviewer: ああそうですか。でも他の会社はそういうことをやっていますか。 #00:39:23-4#

309 Faith2: 今のところまだそこまではないと思います。 #00:39:30-9#

310 Interviewer: そういう計画とかありますか。 #00:39:35-7#

311 Faith2: ただ、あれです。オリンピックがある予定だったので、必ずそういう需要は必要とさ
312 れているとは思っていました。で、計画をされて、まだうちの会社では計画はされていませ
313 ん。 #00:40:00-8#

314 Interviewer: ではね。多分基本的に日本であるリスピークの未来についてどう思いますか。
315 #00:40:18-8#

316 Faith2: でも、そのうちもっとこう AI とか発達してくるとおそらくリスピークではなくて音声
317 直接パソコンとかにつないでやるという可能性はあると思います。今はまだそこまで発達して
318 ないで、人は間に入っているんですけど。いずれはなくなってしまうかなあと考えていま
319 す。 #00:40:53-6#

320 Interviewer: それは、ちょっとリスピーカーのために不安ですか。 #00:40:59-4#

321 Faith2: そうです。でも、生放送の字幕はなくならないので、リアルタイム字幕自体は変わら
322 ないので、そのこの何でしょう。リスピーク自体はなくなっちゃってとしても放送自体はが
323 続くであれば問題はないかなあという感じです。 #00:41:20-1#

324 Interviewer: 多分なんか校正者もひつようですね。 #00:41:27-1#

325 Faith2: そうですね。AI のためやっぱり修正者はまだ残ると思います。 #00:41:31-7#

326 Interviewer: そうですね。今はね。ちょっとストレス とかを話しましたが、その仕事のいいと
327 ころはどれですか。 #00:41:46-1#

328 Faith2: でも、やっぱり必要としてる。耳が不自由とか後は高齢者そういう人とかに見てもら
329 えてるといのは見てる人がいると思うとやっぱりやりがいがありますね。 #00:42:09-1#

330 Interviewer: はい。そしてあの多分楽しい部分もありますね。 #00:42:27-7#

331 Faith2: そうですね。あれでも、テレビが好きだったので、やっぱりその仕事にたたさわれて
332 いるのはうれしいかもしれないです。 #00:42:42-3#

333 Interviewer: そうですね。じゃあの、この仕事を見つけた時は何でしょう普通の就活でしたか。
334 #00:42:55-0#

335 Faith2: そうですね。さっき言った専門学校にそういうお知らせが聞いていて #00:43:08-7#

336 Interviewer: そうですね。 #00:43:11-1#

337 Faith2: ただ自分が入る時はまだリアルタイムがやってなかったんですね。 #00:43:16-5#

338 Interviewer: そうですね。 #00:43:19-6#

339 Faith2: なので、最初もそういうドラマとかさっき言ったバラエティーとかに字幕を付けてい
340 ました。 #00:43:30-2#

341 Interviewer: はい。 #00:43:32-3#

342 Faith2: でそういうを見て入りました。会社に。 #00:43:38-4#

343 Interviewer: では会社は例えば専門学校とかで人員を探していますか。そういうことですか。
344 #00:43:49-6#

345 Faith2r: 専門学校もですし、大学とかもしています。? #00:43:57-3#

346 Faith2: じゃ大学か専門学校に卒業した人例えば別に探していますか。それは必要ですか。卒
347 業した人。 #00:44:16-1#

348 Interviewer: ああ。そうですね。必要だと思いますか。なんと言ったらいかな。大学を卒業
349 した人が #00:44:40-2#

350 Faith2: はい。どなんだろうな。 #00:44:52-2#

351 Interviewer: 大学の学士がある人はお好みですか。会社の人員について。 #00:45:05-1#

352 Faith2: うちの会社は基本的にはなんでしょうね。そういったこと分けてることはしてないで
353 す。 #00:45:16-2#

354 Interviewer: そうですね。 #00:45:19-5#

355 Faith2: 基本的にはリスピークなんでしょうね。多少の知識を持てまれうんですけど。大学を s
356 津行しないとダメとかそういうことではないです。 #00:45:35-6#

357 Interviewer: 分かりました。リスピーカーのみなさん大体どういうバックグラウンドがあり
358 ますか。 #00:45:43-6#

359 Faith2: それこそ自分みたいな専門学校出てる人大学を出た人後は受湯採用ってわかりますか。
360 #00:46:00-9#

361 Interviewer: 分かりません。 #00:46:00-0#

362 Faith2: 別の会社で働いている人、そこをやめて途中からうちに同僚してくれるってことはあ
363 ります。本当にいろんな人はあります。 #00:46:15-1#

364 Interviewer: それは字幕を作っている会社ですか。 #00:46:20-0#

365 Faith2: ないです。全然違うとことからうちに來てる人ともいます。 #00:46:27-6#

366 Interviewer: お好きなバックグラウンドとかがありますか。ありませんか。例えばじゃなリス
367 ととか #00:46:31-5#

368 Faith2: でも、います。新聞記者をやった人がいます。ただ、なんだろうな。その人は確かに
369 政治的なこととか。経済にも詳しいですけど、でも別にわざわざその人が元新聞記者だから取
370 ったこと本当にはないですね。きてみて、面接をしてみて、会話に問題がないとかかそういうと
371 ころをおもにとっていると思います。 #00:47:18-9#

372 Interviewer: そうですね。いろいろなところから來ました人ですか。 #00:47:25-4#

373 Faith2 そうですね。本当にまあ今のところ日本人しかいないんですけど。でも、あまりそうい
374 うバックグラウンドでなんかはじめてしまったりことはあまりないと思いますね。 #00:47:42-1#

375 Interviewer: 多分日本語力が高い人必要は何です。日本人は多いです。 #00:47:49-8#

376 Faith2: 会話というか、正しい漢字を選ぶとか。そういうことをやるのでどうしてもそうです
377 ね。日本人になりがちだと思います。 #00:48:02-5#

378 Interviewer: はい。じゃさっきはちょっとその字幕の対象者例えば何も聞こえない人とかにつ
379 いて話しました。その人たちからフィードバックもらう時がありますか。 #00:48:18-8#

380 Faith2: 直接うちの会社につてことほとんどありません。ただ、そういう耳が聞こえない人と
381 かがこういう字幕だったら見やすいよねみたいなことを目にすることはあります。 #00:48:40-
382 2#

383 Interviewer: そうですね。私の質問は今大体できていますが、他に何がありますか。例えば面
384 白いこと仕事について。私は知らないこととか。 #00:49:05-9#

385 Faith2: でもそうですね。自分は本当にさっき言ってたあそのトラブルとかにも対処できな
386 きゃいけなちゃってすごいパソコンに詳しくなりました。全然最初はパソコンにさわってこな
387 かったですけど、今はパソコンのシステムとか、そういうことにもかかわっています。 #00:49:52-
388 0#

389 Interviewer: それも大切なことですね。 #00:49:55-5#

390 Faith2: そうですね。 #00:50:02-0#

391 Interviewer: ああ、もひとつがあります、そのさっき言った練習と教育どのぐらいかかります
392 か。 #00:50:08-5#

393 Faith2: 二か月ぐらいはやっぱり最低限でも、リスピークを練習する必要があります。 #00:50:24-
394 1#

395 Interviewer: それは短いとおもいますね、二か月だけ。 #00:50:29-4#

396 Faith2: そうですね。対応する、われわれは担当する番組によってもレベルがあるですね。さ
397 っきも言ってたバラエティーだと難しいとかそういうがあるので一応二か月やればその一番
398 下のレベルのものには対応できるという感じです。 #00:50:58-0#

399 Interviewer: そうです。何が一番低いレベルの番組ですか。 #00:51:04-0#

400 Faith2: 例えば、ニュースなんですけど、 #00:51:10-4#

401 Interviewer: ニュースですか。 #00:51:12-0#

402 Faith2: ニュースなんですけど、原稿があらかじめできてるもの #00:51:16-3#

403 Interviewer: そうですね。 #00:51:18-2#

404 Faith2: それをほとんど読むだけみたいな時もあります。書いてあるものをほとんど繰り返し
405 読めばやってみたい時があるので、そういう時は多分それならさっき言った二か月練習すれば
406 いいみたいな。それならできると思います。 #00:51:40-9#

407 Interviewer: でも、時々自分の制度率字幕はどのぐらいできるとかについてフィードバックが
408 もらいますか。 #00:51:55-4#

409 Faith2: それはほぼ毎回フィードバックしています。 #00:51:57-6#

410 Interviewer: 毎回そうですね。それは誰からのフィードバック。 #00:52:00-8#

411 Faith2: 自らですかね。自分たちで放送後にもしくはトレーニングが終わった後にリスピーク
412 のその精度どのぐらい正しくあったかっていうのを見ることができます。 #00:52:24-7#

413 Interviewer: じゃあの何パーセントぐらいの精度が大丈夫ですか。そういう基準は。 #00:52:31-
414 9#

415 Faith2: 一応目指してるのは80パーセント以上です。 #00:52:36-5#

416 Interviewer: そうです。それは最初ですか。 #00:52:44-0#

417 Faith2: ただ難しいのはどこまで合ってるかというのはその大丈夫かという判断するのは難しく
418 て #00:52:58-5#

- 419 Interviewer: 確かに #00:53:00-4#
- 420 Faith2 例えば日本語だと漢字によって意味が変わっちゃいますよね。 #00:53:04-1#
- 421 Interviewer: はい。 #00:53:06-6#
- 422 Faith2: でも、音は同じじゃないですか。 #00:53:11-7#
- 423 Interviewer: そうですね。 #00:53:14-1#
- 424 Faith2: なので、あくまでもリスピーカーの制度率導く出す場合は音通り話してればオッケっ
425 てことにしてます。 #00:53:26-3#
- 426 Interviewer: そうですね。ほかは修正者のこと何ですね。 #00:53:30-4#
- 427 Faith2: そうです。なので間違ったまま置くってしまった場合はそれは修正者のミスになりま
428 す。 #00:53:40-2#
- 429 Interviewer: わかりました。でも、トレーニングのあとにもこのフィードバック毎回やってま
430 すか。 #00:53:51-9#
- 431 Faith2 そうですね。
- 432 Interviewer: 私の質問は今終われいましてので、ありがとうございます、本当に。

Anhang 3 Zielgruppeninterviews

Interview T1

- 1 1. お名前 廣川麻子
- 2 2. 性別 女性
- 3 3. 年齢 47 歳
- 4 4. 職業 大学職員、NPO 団体代表
- 5 5. 難聴のレベルは何ですか。 両側感音性難聴 2 級（両耳とも 100dB）
- 6 6. 何歳のころに聴力を失いましたか。 先天性です。（生まれた時から聴覚障害者です）
- 7 7. 人工内耳か補聴器を使用していますか。使用すると聴力はどうなりますか。
8 両耳に補聴器を装用しています。補聴器装用時の聴力は計測しておりませんが、環境
9 音（電話の音、救急車などの緊急車両の音など）、音楽の音は聞こえますが、言葉を
10 含め、正確な判別は出来ません。
- 11 8. あなたの母語は何語ですか。例えば、日本手話か、日本語か、もしくはどちらも母語だ
12 と認識していますか。
13 どちらも母語だと認識しています。
- 14 9. テレビを見る時にどのくらいの頻度で字幕を使用していますか。
15 例えば、一か月〇〇回、一週間〇〇回、毎日...

- 16 テレビを視聴するときは100%字幕を設定しています。
- 17 10. どの種類の番組で字幕を使用しますか。字幕が必要ではない種類の番組もありますか。
- 18 字幕がない番組は視聴しません。
- 19 11. 生放送でも字幕を使用しますか。(例えばニュースやスポーツやほかの事前に字幕を作る
- 20 ことが出来ない放送)
- 21 はい。上記回答に重なりますが、字幕がつくものは全て字幕を設定して視聴します。
- 22 12. 日本のテレビの字幕の量についてどう思いますか。
- 23 足りですか。足りませんか。
- 24 足りません。事前制作の番組にはほぼ付与されるようになりましたが、CMにはほぼつ
- 25 きません。またリアルタイム番組（特にバラエティー情報番組）にはつかないことが
- 26 多いです。
- 27 13. 日本のテレビの字幕全般の質についてどう思いますか。どのように改善できますか。
- 28 ここ数年の間に、質は大変高くなっていると感じています。
- 29 14. リアルタイム字幕全般の質についてどう思いますか。どのように改善できますか。
- 30 現在の技術では2秒ほど遅れるとのことですが、大変ストレスです。特にニュースな
- 31 どでは画面とズレているため、理解が難しくなります。
- 32 15. 字幕を見る時、何か問題はありますか。どのような問題ですか。
- 33 テレビの機種によって、字幕を画面に被せたり、外に出したり（イン、アウト）する
- 34 機能があったり、なかったりします。
- 35 16. 字幕について何か変えてほしいことはありますか。どういうことですか。
- 36 完全字幕、要約字幕と選択できると良いです。
- 37 17. 良い字幕を作るために大切なことはどんなことですか。
- 38 音を文字化することで、雰囲気伝えることが大切。音の説明を詳しくしてほしい場
- 39 合があります。
- 40 18. 日本のテレビ局を比較して、ほかより良い字幕を製作しているテレビ局はあると思いま
- 41 すか。それはなぜですか。
- 42 そのような視点で、見たことがありませんので、分かりません。
- 43 19. 日本のテレビ局を比較して、ほかより良いリアルタイム字幕を製作しているテレビ局が
- 44 あると思いますか。どうしてですか。
- 45 そのような視点で、見たことがありませんので、分かりません。
- 46 20. 日本のテレビで出る字幕は読みやすいと思いますか。読みにくいと思いますか。それは
- 47 なぜですか。

48 現在の字幕は読みやすいと思っています。

49 21. 字幕で話者の交代を表す時どうすれば良いと思いますか。

50 話者名をつけると良いです。(山田)のように。

51 22. 字幕の速さについてどう思いますか。速すぎますか。遅すぎますか。ちょうど良いです

52 か。

53 話されていることと同じスピードが良いです。臨場感が大切です。早口なら、字幕の

54 表示も早くあるべきです。

55 23. 話されたことが全部書いている字幕と、修正したり略したりしている字幕のどちらが良

56 いとおもいますか。話されたことが全部書いてある字幕が良いと思う場合、もしその方

57 が修正の入っている字幕よりも間違いが多いとしても、そちらが良いと思いますか。

58 私個人的には、完全字幕を希望します。ただし、状況によります。スピードを優先す

59 るか、正確を優先するかは番組の性質によると考えます。

60 24. リアルタイム字幕がどのように製作されているということをご存じですか。どのように

61 製作されていると思いますか。

62 音声認識ソフトなどを使って修正する方式、手打ち方式とあると承知しています。

63 25. リアルタイム字幕を製作する時にどのような問題があると思いますか。

64 上記にも回答しましたが、秒遅れの問題が大きいです。

65 26. あなたの考える、テレビを聴覚障害者にとってバリアフリーにするための一番良い方法

66 は何ですか。

67 制作資金の問題。キー局はだいぶ進んだと思いますが、地方局は制作資金が不足しているた

68 め、字幕付与が進んでいないと聞いています。国の補助をもっと強化するべきです。また、

69 字幕制作会社をもっと増やす必要があります。

70 27. ほかに付け加えたいことはありますか。

71 テレビの問題もありますが、オンライン動画への字幕付与も重要と考えています。

Interview T2

- 1 1. お名前 瀬川奈美
- 2 2. 性別 女性
- 3 3. 年齢 45 歳
- 4 4. 職業 会社員
- 5 5. 難聴のレベルは何ですか。 右 90dB, 左 100dB
- 6 6. 何歳のころに聴力を失いましたか。生まれつき聞こえない
- 7 7. 人工内耳か補聴器を使用していますか。使用すると聴力はどうなりますか。

8 →現職場で働き始めた頃から、(26 歳～現在)補聴器・人工内耳どちらも使用せず。

9 8. あなたの母語は何語ですか。例えば、日本手話か、日本語か、もしくはどちらも母語だ

10 と認識していますか。

11 →手話が母語だと認識しています

12 9. テレビを見る時にどのくらいの頻度で字幕を使用していますか。例えば、一か月〇〇回、

13 一週間〇〇回、毎日...

14 毎日使用する。

15 10. どの種類の番組で字幕を使用しますか。字幕が必要ではない種類の番組もありますか。

16 全ての番組で字幕を使用する。

17 11. 生放送でも字幕を使用しますか。(例えばニュースやスポーツやほかの事前に字幕を作る

18 ことが出来ない放送)

19 使用する。

20 12. 日本のテレビの字幕の量についてどう思いますか。足りますか。足りませんか。

21 足りないと思う。

22 13. 日本のテレビの字幕全般の質についてどう思いますか。どのように改善できますか。

23 質は高いと思う。

24 14. リアルタイム字幕全般の質についてどう思いますか。どのように改善できますか。

25 質は高い。漢字の誤変換があってもさほど気にならない。

26 15. 字幕を見る時、何か問題はありますか。どのような問題ですか。

27 見慣れているので、特に問題はない。

28 16. 字幕について何か変えてほしいことはありますか。どういうことですか。

29 特にない

30 17. 良い字幕を作るために大切なことはどんなことですか。

31 当事者を交えて改善していくこと。

32 18. 日本のテレビ局を比較して、ほかより良い字幕を製作しているテレビ局はありますか。思いま

33 すか。それはなぜですか。

34 標準化しているので、特になし。

35 19. 日本のテレビ局を比較して、ほかより良いリアルタイム字幕を製作しているテレビ局が

36 あると思いますか。どうしてですか。

37 上記 10) に同じ。

38 20. 日本のテレビで出る字幕は読みやすいと思いますか。読みにくいと思いますか。それは

39 なぜですか。

- 40 読みやすいと思う。
- 41 21. 字幕で話者の交代を表す時どうすれば良いと思いますか。
- 42 字幕の色でわかるようにするとよい。
- 43 22. 字幕の速さについてどう思いますか。速すぎますか。遅すぎますか。ちょうど良いです
- 44 か。
- 45 ちょうどよい。
- 46 23. 話されたことが全部書いている字幕と、修正したり略したりしている字幕のどちらが良
- 47 いとおもいますか。話されたことが全部書いてある字幕が良いと思う場合、もしその方
- 48 が修正の入っている字幕よりも間違いが多いとしても、そちらが良いと思いますか。
- 49 間違いが多かったとしても、話されたことが全部書いてある字幕がよい。
- 50 24. リアルタイム字幕がどのように製作されているということをご存じですか。どのように
- 51 製作されていると思いますか。
- 52 今はリスピーク型。いずれ音声認識型になるだろうと考えている。
- 53 25. リアルタイム字幕を製作する時にどのような問題があると思いますか。
- 54 誤変換の対策。
- 55 26. あなたの考える、テレビを聴覚障害者にとってバリアフリーにするための一番良い方法
- 56 は何ですか。
- 57 公共の場（駅、病院、役所など）のテレビで字幕を出すこと。
- 58 27. ほかに付け加えたいことはありますか。
- 59 特になし

Interview T3

- 1 1. お名前 藤川 太郎
2 2. 性別 男
3 3. 年齢 41
4 4. 職業 団体職員
5 5. 難聴のレベルは何ですか。 100db 以上
6 6. 何歳のころに聴力を失いましたか。 先天性
7 7. 人工内耳か補聴器を使用していますか。使用すると聴力はどうなりますか。
8 補聴器使用。音の判別はできるが、言葉としては理解できない。
9 8. あなたの母語は何語ですか。例えば、日本手話か、日本語か、もしくはどちらも母語だと認識していますか。
10 書記日本語、日本手話言語の両方が母語である
11 テレビを見る時にどのくらいの頻度で字幕を使用していますか。
12 例えば、一か月〇〇回、一週間〇〇回、毎日...
13 ほとんど見ない
14 9. どの種類の番組で字幕を使用しますか。字幕が必要ではない種類の番組もありますか。
15 全ての番組で字幕を使用する。字幕が必要でない種類の番組はない。
16 10. 生放送でも字幕を使用しますか。(例えばニュースやスポーツやほかの事前に字幕を作る
17 ことが出来ない放送)
18 はい、字幕が付いていれば見ます。字幕がなければ見ません。
19 11. 日本のテレビの字幕の量についてどう思いますか。
20 足りですか。足りませんか。
21 字幕が付いている番組が増えましたが、このまま全ての番組に字幕が付いてほしいです。
22 12. 日本のテレビの字幕全般の質についてどう思いますか。どのように改善できますか。
23 時々、字幕が遅れる時があるので、リアルタイム（実際の話と同じタイミングで字幕が表示される）で理解できれば良い。最近はテロップも増えており、テロップはリアルタイムで付加されているのでわかりやすいものが多い。なお、日本の場合は、誤字は少ないです。
24 13. リアルタイム字幕全般の質についてどう思いますか。どのように改善できますか。
25 ときどき字幕が遅れてしまうのが残念。改善方法はわかりません。
26 14. 字幕を見る時、何か問題はありますか。どのような問題ですか。
27
28
29
30
31

32 字幕が遅れる時がある以外は特に問題は感じていない。

33 15. 字幕について何か変えてほしいことはありますか。どうのことですか。

34 字幕が遅れる時がある。それ以外は特に問題は感じていない。

35 16. 良い字幕を作るために大切なことはどんなことですか。

36 AIではなく今まで通り人手での入力をお願いしたい。

37 17. 日本のテレビ局を比較して、ほかより良い字幕を製作しているテレビ局はあると思いま

38 すか。それはなぜですか。

39 大手のテレビ局の字幕には満足しているが、小さなテレビ局では字幕が付かないこと

40 が多い。字幕の質より量（字幕が付いている番組）を増やしてほしい。

41 18. 日本のテレビ局を比較して、ほかより良いリアルタイム字幕を製作しているテレビ局が

42 あると思いますか。どうしてですか。

43 小さなテレビ局はお金がないからだと思っている。

44 19. 日本のテレビで出る字幕は読みやすいと思いますか。読みにくいと思いますか。それは

45 なぜですか。

46 読みやすい。背景と重ならないようにするために、字幕が表示される位置が、随時変

47 更されるため。

48 20. 字幕で話者の交代を表す時どうすれば良いと思いますか。

49 日本の場合は、話者によって字幕の色を変えたり、話者が分かるように話者の名前が

50 字幕に入るため、わかりやすい。

51 21. 字幕の速さについてどう思いますか。速すぎますか。遅すぎますか。ちょうど良いです

52 か。

53 ちょうどよい。

54 22. 話されたことが全部書いている字幕と、修正したり略したりしている字幕のどちらが良

55 いとおもいますか。話されたことが全部書いてある字幕が良いと思う場合、もしその方

56 が修正の入っている字幕よりも間違いが多いとしても、そちらが良いと思いますか。

57 誤字があったとしても話されたことは全部書いてもらいたいです。

58 23. リアルタイム字幕がどのように製作されているということをご存じですか。どのように

59 製作されていると思うますか。

60 はい、どのように制作されているか知っています。

61 24. リアルタイム字幕を製作する時にどのような問題があると思いますか。

62 字幕製作が追い付かなくなる時がある為、（字幕製作者はリアルタイムで視聴し）、

63 実際の放送は数秒送らせて放映する方法が良いと思います。

64 25. あなたの考える、テレビを聴覚障害者にとってバリアフリーにするための一番良い方法
65 は何ですか。

66 手話言語通訳者をつけることです。（遅れてしまう字幕と違ってリアルタイムで情報
67 を把握することができます。）

68 26. ほかに付け加えたいことはありますか。

69 無し

Interviews T4 bis T8

1. 名前	T4	T5	T6	T7	T8
2. 性別	女性	男性	男性	男性	女性
3. 年齢	44	37	34	58	36
4. 職業	事務・オフィス系	IT・エンジニア系	IT・エンジニア系	事務・オフィス系	事務・オフィス系
5. 難聴のレベル	重度難聴 (Profound) 100dB 以上	重度難聴 (Profound) 100dB 以上	重度難聴 (Profound) 100dB 以上	重度難聴 (Profound) 90dB 以上	中高度難聴 (Moderately severe) 56 ~ 69dB 補聴器がないと聞き取りが難しい。
6. 聞こえなくなったのは何歳のときですか (年齢)	3	生まれつき	生まれつき	4	4
7. 補聴器等は使っていますか	補聴器 (片耳) 人工内耳 (片耳)	補聴器 (両耳)	人工内耳 (両耳)	補聴器 (両耳)	補聴器 (片耳)

8. 補聴器・人工内耳を装用すると聴力はどうなりますか。	人工内耳：30db 補聴器：60～70db	音としては入るが、言葉としては入らない	20～30dB	聴力は95dBSPL → 45dBSPLに改善される。だが語音弁別能は変わらない。	30dB くらいになる
9. 第一言語または母語	日本語	日本語	日本語	日本語	日本語
10. コミュニケーション方法	口話; 読話; 日本語対応手話; 筆記; 音声認識	口話; 読話; 日本語対応手話; 筆記	口話; 読話; 筆記	口話; 読話; 日本手話; 日本語対応手話; 筆記; 音声認識	口話; 読話; 日本語対応手話
11. どのようなテレビ番組をよく視聴していますか	ニュース; パラエティ・情報・ワイドショー; ドラマ・映画; 音楽; 趣味・学習; ドキュメンタリー	ニュース; スポーツ; ドキュメンタリー	ニュース; パラエティ・情報・ワイドショー; ドラマ・映画; ドキュメンタリー	ニュース; アニメ; パラエティ・情報・ワイドショー; スポーツ; ドラマ・映画; ドキュメンタリー	バラエティ・情報・ワイドショー; ドキュメンタリー
12. テレビ番組をみるとき、字幕をつけていますか	必ずつける	必ずつける	必ずつける	必ずつける	つけたりつけなかったりする

13. 生放送でリアルタイム字幕をつけますか	必ずつける	必ずつける	字幕はつけない	必ずつける	つけたりつけなかったりする
14. テレビをみるとき、補聴器や人工内耳等をつけていますか	つけたりつけなかったりする	つけたりつけなかったりする	必ずつける	つけたりつけなかったりする	必ずつける
15. 楽しみにしている番組をみるとき、音声はつけますか	必ず音をだしている	音声はつけたりつけなかったりする	必ず音をだしている	音声はつけたりつけなかったりする	必ず音をだしている
16. 字幕がなくてもいい番組というのがありますか。	ない	ない	ある	字幕がないから気にならない（見ようと思わない）番組は多数ある。	ない
17. 数少ない手話番組はみていますか？	ときどき見る	ほとんど見ない、興味がな	ときどき見る	ときどき見る	ほとんど見ない、興味がな

18. 字幕がついているテレビ番組の量について、どう思いますか。足りませんか。足りませんか。	足りません。BS や CS を含め、すべての番組に字幕がついているのが希望です。	足りていない	理想は１００％字幕付与	直近の 20 年でかなり字幕がつくようになった。だが緊急災害時の生放送は命に関わる重要な情報だが、ほぼ字幕がつかない。これは字幕送出のシステム上の問題があるためである。テレビの政見放送も公職選挙法の制限が上位にあるため、字幕を付けられない政見放送(参議院議員選挙選挙区選挙、都道府県知事選挙)がある。いずれも当事者団体が長年に渡って改善要望をし続けており、少しずつ付与されるようになってきている。	足りない
---	---	---------------	--------------------	---	-------------

19. 字幕がついてなくて不満を感じている番組があれば教えてください	語学の教育番組、国会中継、地方のテレビ番組、深夜のバラエティ番組など	早朝や深夜の番組。 例えばNHKニュースは7時台には字幕がつくが6時台につかない	生放送番組における字幕付与（遅延や字幕がないケース）	早朝（午前7時以前）のニュース番組。国の指針では午前6時以前は対象外になっている。	
20. テレビの字幕の質についてどう思いますか。どのように改善できると思いますか	おおむね良いと思う。ときどき字幕がかぶって大事な場面が見えなくなるので、文字の大きさや速さを変えられるといいと思う。	特に不満はない	テレビ局によって字幕の見やすさが異なるため、更なる標準化を目指した方が良いのではないか	かなり満足している。	文字の位置・大きさ（もともとついてるものに干渉したり、画面を邪魔しないもの）にもう少し思い至ればいいなと思う

21. リアルタイム字幕の質についてどう思いますか。どのように改善できると思いますか	技術的に仕方がないが、表示が遅い。放送時間を今より4秒遅らせる。AIが字幕入力ができるようになる。	リアルタイム字幕は若干遅れて入るが、CMに入るとその直前の部分の字幕が表示されない。これも表示するようにしたい。	映像と字幕の不一致による視聴ストレス。放送時間を遅らせてもいいので、一致したコンテンツを放送してほしい	誤字が散見されるのがやや不満。画面の内容から約8秒以上遅れて字幕が表示される。画像送出を遅らせて、字幕ができてから画面と一緒に送出すればよいと思う。	アナウンサーの原稿を渡す？
22. 字幕を見る時、何か問題がありますか。どのような問題ですか	ときどき字幕がかぶって大事な場面が見えなくなること。	上参照	生放送番組における字幕付与、字幕付与の標準化等	生放送の字幕は、CMに入るときや場面が変わるときなど、字幕のしっぽをちょん切ってしまうので、話の内容を最後まで完全に理解することができない。	特になし

23. 字幕について何か変えてほしいことはありますか。どのようなことですか。	字幕のマークをもっと活用してほしい。ユーザーと字幕製作者で交流し、ユーザーの声をフィードバックしてほしい。	上参照	上記の通り	盲ろう者にも扱えるように、テキスト送信してほしい。今の放送の規格では、コードが異なるためテキスト化ができない。	
24. よい字幕を作るために大切なことはどんなことだと思いますか	字幕は聞こえる人にも便利であること	正確さと表示までの速度	ユーザーレビュー	字幕を作る側にできるだけ当事者が関わること。	字幕を作成する人への支援
25. テレビ局を比較して、よい字幕を製作しているテレビ局はどこだと思いますか。それはなぜですか	① NHK 文字の配置が上手い。 ② BS 朝日衛星放送のなかでは字幕付与の努力を感じる	分からない	NHK。見やすさ、話者毎の字幕の色、タイミング、精度	キー局と言われる主要局は字幕が多い。逆にローカル局は財政上の問題があり、字幕がつかないことが多い。視聴者のために国の支援が必要だと思う。	

26. テレビ局を比較して、よいリアルタイム字幕を製作しているテレビ局はありますか。それはなぜですか	NHK 生字幕なのに色分けしたり、リスピーク方式を取り入れたり、様々な企業努力をしているから。	分からない	上記の通り	NHK。現在音声認識を活用して、ローカル局でも災害時でもリアルタイムで字幕送出できるシステムのテストをしている。	
27. 日本のテレビ字幕は読みやすいと思いますか。読みにくいと思いますか。理由も教えてください	読みやすいです。登場人物ごとに色分けされており、話者の位置に合わせて文字が配置され、読みやすいフォントを使っています。	だいたい読みやすいが、たまに画面で必要な部分の上に字幕が表示されるときがある。その時はちょっと読みにくい。	英語字幕と比較すると、比較的に読みやすいと感じる	欧米の一部の字幕は流れるように動く(スクロール)ことがあるが、日本の字幕は固定表示で視線移動が少なく、読みやすいと思う。	

28. リアルタイム字幕で話者の交代を表すとき、どうすればよいと思いますか。	色分けする、名前を入れる	色、フォントを変える。	字幕の色を変更する等	サン/おわりだ。アシタカ/おわらない。私たちがいる。のように、話者の表示があるとよい。	色分け
29. 字幕の早さについてどう思いますか。	ちょうどいい	ちょうどいい	ちょうどいい	早くて読みきれないことがある	遅いと思う
30. 表示が遅いと感じても話されたことが全部書いている字幕と、早いが誤字があったり要約された字幕のどちらがよいと思いますか	表示は遅くなるが、全部かいてある字幕	表示は遅くなるが、全部かいてある字幕	早いけど、誤字や要約されている字幕	早いけど、誤字や要約されている字幕	表示は遅くなるが、全部かいてある字幕

31. リアルタイム字幕がどのように製作されているかご存じですか。どのように製作されていると思いますか。	リアルタイム字幕の現場に見学に行ったことがあります。3名で入力し、1名が確認作業を行うという4名1チームで制作していました。大変な集中力で作業していました。	リスピークなどは知ってる	はい。字幕制作会社によって制作される	スピードワーク方式、パケツ(事前登録活用)方式、リスピーク方式などの方法を使っている。	不明
--	--	--------------	--------------------	---	----

32. リアルタイム字幕の問題はどれだとおもいますか（複数可）	表示が遅れるため、画面と字幕の話題がずれてしまうこと；番組が終わると字幕が途中で切られてしまうこと；音情報（環境音の説明など）が抜けてしまうこと；リアルタイム字幕がつかない生放送があること；外国語のリアルタイム字幕がないこと	番組が終わると字幕が途中で切れてしまうこと；リアルタイム字幕がつかない生放送があること	表示が遅れるため、画面と字幕の話題がずれてしまうこと	表示が遅れるため、画面と字幕の話題がずれてしまうこと；番組が終わると字幕が途中で切られてしまうこと；音情報（環境音の説明など）が抜けてしまうこと；リアルタイム字幕がつかない生放送があること；話者が誰なのかわかりにくくなる	表示が遅れるため、画面と字幕の話題がずれてしまうこと；番組が終わると字幕が途中で切られてしまうこと
--	---	--	-----------------------------------	---	--

33. あなたの考える、聴覚障害者にとってテレビをバリアフリーにするために一番良い方法は何だと思いますか	いろいろな人が一緒に楽しめる状態とはどんなことを考えること	字幕と音声の同期。生放送でも画像と音声を多少遅延させればいいのか？	Netflix のように、すべてのコンテンツに対して字幕付与がスタンダードであり、言語も自由に選択できること。バリアフリーの観点ではなく、ユニバーサルデザインの概念で字幕付与普及を進めてほしい	H.702 の規格活用	
--	-------------------------------	-----------------------------------	--	-------------	--

34. 日本の字幕について、思うことがあれば教えてください	高齢社会ですが、字幕の仕方がわからない人が多いようです。聞こえていても、赤ちゃんや病人や在宅ワークをしている家族のいる人も使うそうです。また、洗い物をしているときに水音で音声が聞こえにくくなるので、そういうときに字幕を使う人もいます。 現在、日本は6時から25時		上記の通り	テレビCM部分はシステム上、字幕をつけにくかった。改善の動きがあるがまだまだである。広告会社は聴覚障害の消費者をみすみす見逃している。これまで当事者不在でシステムを作ってきた弊害がここにも出ている。	
--------------------------------------	---	--	--------------	--	--

	までが 字 幕 付 与 の 対 象とな ってい ます。 また、 大 勢 の人が 話 す 生 放 送 も 字 幕 をつけ なくて もいい という ことが あり、 まだ字 幕がな くて残 念 に 思うこ とがあ り ま す 。 コマー シャル や 国 会 中 継 も 字 幕 がない ため政 治 参 加がで きませ ん 。 インタ ーネッ ト 番 組 も 字 幕 がつい ている ものは 少ない です。				
--	--	--	--	--	--

	それら があまり 問題にあ がって こない ので、 残念 に思っ ていま す。				
--	--	--	--	--	--

Kurzfassungen

Kurzfassung Deutsch

In letzter Zeit ist das gesellschaftliche Bewusstsein für die Wichtigkeit des barrierefreien Zugangs zu Medien für Menschen mit Behinderung stetig gestiegen. Eine Art, diesen Zugang zu ermöglichen, ist die Nutzung von Untertitelung für schwerhörige und gehörlose Personen. Besonders schwierig zu gewährleisten ist diese Untertitelung bei Live-Sendungen, wenn keine Untertitel im Voraus angefertigt werden können. Daher befassen sich auch immer mehr wissenschaftliche Arbeiten im Bereich der audiovisuellen Translation mit der Live-Untertitelung. Ziel dieser Masterarbeit ist es, einen Überblick über die momentane Situation der Live-Untertitelung in Japan zu geben und dabei besonders auf die Verwendung der Respeaking-Methode einzugehen. Hierzu wurde zunächst die japanischsprachige Literatur zu diesem Forschungsgebiet aufgearbeitet. Anschließend folgten qualitative Interviews mit Vertretern des japanischen öffentlichen Fernsehsenders NHK. Um ein ganzheitlicheres Bild zu bekommen, befasst sich die Arbeit auch mit der Hauptzielgruppe der Untertitelung, jener der gehörlosen und schwerhörigen Personen. Die Interviews mit diesen TeilnehmerInnen wurden schriftlich abgewickelt. Die zugrundeliegende Methodik dieser Befragungen und deren Analyse ist die qualitative Inhaltsanalyse nach Mayring. Diese Untersuchung brachte folgende Ergebnisse hervor. Bei NHK findet seit Anfang der 2000er Jahre Live-Untertitelung statt. Diese wird durch verschiedene Methoden, zum Beispiel Steno-Tastaturen, aber auch mit direkter Spracherkennung und Respeaking bereitgestellt. Auffallend ist, dass die Arbeit im Team stattfindet, wobei die Kompetenzen auf RespeakerInnen und EditorInnen verteilt sind. Es ist keine spezielle Vorbildung verlangt, und die Ausbildung findet größtenteils in-house statt. Für die Qualitätskontrolle gibt es von staatlicher Seite keine Vorgaben, die über die auf jeden Fall zu untertitelnden Sendezeiten hinausgehen. Eine der größten Herausforderungen bei der Live-Untertitelung ist die große Häufigkeit von Homonymen in der japanischen Sprache. Die gehörlosen und schwerhörigen Befragten stufen die Qualität der Untertitelung generell als hoch ein. Problematisch sei vor allem die Verzögerung und die Beschränkung der Untertitelung auf bestimmte Tageszeiten.

Kurzfassung Englisch

Lately there has been increasing awareness in society of the importance of media accessibility for people with disabilities. One way to ensure this access is the use of Subtitling for the Deaf and Hard of Hearing (SDH). This type of subtitling is especially challenging to provide for live broadcasts when subtitles cannot be produced in advance. For this reason, more and more papers within the field of audiovisual translation are covering this issue. This thesis aims to provide an overview of the current situation of live subtitling in Japan, focusing on the respeaking method. Initially, an analysis of the existing literature on live subtitling available in Japanese language was carried out. Subsequently, qualitative interviews were conducted with representatives of the Japanese public broadcaster NHK. In order to get a holistic perspective on the issue, the paper also takes into consideration the main target audience of deaf and hard of hearing individuals. Interviews with this group were conducted in written format. The methodology used for conducting interviews and analysing the data is the Qualitative Content Analysis according to Mayring. The research brought about the following results: NHK has been providing live subtitling since the early 2000s, using a variety of methods such as stenography keyboards, respeaking and direct voice recognition. Notably, the subtitling is carried out as a team, with competencies being distributed between respeakers and editors. There are no specific educational requirements for respeakers and training is carried out in-house. There are no governmental guidelines for quality control aside from defining the times of day where subtitling has to be provided. One of the major challenges for live subtitling is the prevalence of homonyms in the Japanese language. The deaf and hard of hearing study participants evaluated the quality of the subtitles as generally high. Problem areas are the delay and the limitation of subtitles to certain times of the day.