

Elisabeth HILDEBRANDT

Sprachverständlichkeit im Fernsehen

Vorstellung von ausgewählten Teilaspekten zu diesem Thema
im Kontext der Entwicklung einer Production Guideline

DIPLOMARBEIT

zur Erlangung des akademischen Grades Magistra Artium

Studium: **Tonmeister**

Institut für Komposition und Elektroakustik
Universität für Musik und darstellende Kunst Wien

Betreuer: Mag.art. Dr.rer.nat. Peter Mechtler

Wien 2014

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Diplomarbeit selbständig angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder in gleicher noch in ähnlicher Form einer anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Wien, 03.03.2014

Vorwort

Besonderer Dank gilt:

Michael Eberhard und **Erich Ebert** vom Südwestrundfunk Stuttgart, die diese Arbeit überhaupt erst ermöglicht haben und mir sämtliche Informationen, Tonstudios und die Hilfe der Kollegen zur Verfügung gestellt haben.

Dr. Peter Mechtler, der diese Arbeit mit großem Einsatz betreut hat.

Marianne Blumers (Medienforscherin des SWR Stuttgart), die mir sehr oft mit Rat und Tat zur Seite stand und mich immer wieder motivierte.

Matthias Eberhardt (SWR), **Elmar Hergenröder** (BR), **Daniel Matejka** (ZDF), **Max Kiefer** (WDR) und **Askan Siegfried** (NDR), die mich mit der AG Audio der Fernsehbetriebsleiter sehr unterstützt haben und mir viele Informationen geliefert haben.

Josef Lipp (NDR), der mich vier Tage zum NDR nach Hamburg eingeladen hat, um Versuche im Mischlabor mit Askan Siegfried durchzuführen.

Dr. Dieter Walker und seiner Arzthelferin **Heidrun Bosch**, die den Hörtest aller Probanden durchgeführt haben.

Benedikt David, **Kathrin Dick**, **Claudia Ebert**, **Erich Ebert**, **Bernd Junior Hildebrandt**, **Dr. Alexander Hildebrandt** und **Bärbel Rometsch**, die diese Arbeit unermüdlich Korrektur gelesen haben.

Meinen **21 freiwilligen Probanden**, die den Hörvergleich mit mir durchgeführt haben und dafür zwei Stunden im Funkhaus Stuttgart verweilten.

Christian Simon, der mir oft telefonischen Beistand geleistet hat.

Meinem Verlobten **Markus Bodenseh**, der mir in allen Dingen den Rücken frei gehalten hat, um diese Arbeit abzuschließen und mich in schwierigen Phasen immer wieder zum Lachen gebracht hat.

1 Einleitung	1
2 Grundlagen	4
2.1 Sprechwissenschaften	4
2.1.1 Ästhetische Kommunikation	4
2.1.2 Sprechbildung	7
2.1.3 Sprecherziehung der Schauspieler	12
2.2 Auditive Wahrnehmung	14
2.2.1 Hörfläche des menschlichen Gehörs	15
2.2.2 Frequenzspektrum der Sprache	17
2.2.3 Formanten	19
2.2.4 Verdeckung	20
2.2.5 Altersschwerhörigkeit	24
2.2.6 Bild/Ton Beziehung	27
3 Sprachverständlichkeit im Fernsehen	29
3.1 Zuschauerbeschwerden	29
3.1.1 Zuschauerredaktionen	29
3.1.2 Analyse der Beschwerden	30
3.1.3 Ergebnisse	32
3.2 Hörvergleich	34
3.2.1 Versuchsbeschreibung	35
3.2.2 Probanden	35
3.2.3 Filmausschnitte	36
3.2.4 Herstellung der Mischvarianten	37
3.2.5. Fragebogen	41
3.2.6 Studio „13“ des SWR	42
3.2.7. Versuchsablauf	43
3.2.8 Ergebnisse und Diskussion	45
3.2.9 Lösungsansatz für die Mischung	49

3.3 Abhörsituation während der Produktion	50
3.3.1 Produktionsregie	50
3.3.2 Synchronstudio/NDR - Problematik	51
3.3.3 Empfehlung für Produktionsregionen	53
4 Zusammenfassung und Ausblick	54
5 Literaturverzeichnis	57
6 Anhang	61

Abbildungsverzeichnis

Seite

1	Zusammenwirken der Funktionskreise	8
2	Schematische Darstellung der Erzeugung von Stimmanschall	9
3	Das Ohr mit seinen drei Teilen	15
4	Hörschwelle und Kurven gleicher Lautstärke	16
5	Frequenzbereich von Sprache und Musik	18
6	Schematische Darstellung des Stimmfeldes	18
7	Hörfeld	19
8	Formanten	20
9	Mithörschwelle	21
10	Maskierung, kritische Bandbreite	22
11	Entdeckungs- und Verständlichkeitsschwelle	22
12	Mithörschwelle verursacht durch einen 200 Hz Trompeten-Ton	23
13	Mithörschwellen eines Tones durch Violin-Klänge	23
14	Hörfläche des menschlichen Gehörs	24
15	Schema zum Prinzip des Recruitments	25
16	Altersabhängiger Hörverlust	26
17	Aufteilung der Zuschauerbeschwerden	32
18	Aufteilung in unterschiedliche Genres	33
19	Izotope Show der Naturwunder	40
20	Izotope Frequenzen für Equalizer eines „Show der Naturwunder“ Ausschnitts .	40
21	Grundriss Studio 13 des SWR Stuttgart	42
22	Frequenzgang Studio 13 gemessen mit dem „Acoustilyzer AL1“	42
23	RT60 Studio 13 gemessen mit dem „Acoustilyzer AL1“	43
24	Probanden in Studio 13 vor dem Fernseher (eigenes Foto)	43
25	Vergleich Frage 2 und 3	46
26	Bewertung der Mischvarianten 7 und 10 LU	47
27	Aufteilung der Probanden in unter 50 und über 50 Jahre	48
28	Synchronstudio 1 NDR (eigene Grafik)	51

Tabellenverzeichnis

Seite

1	Altersgruppen für den Hörvergleich	36
2	Ausschnitt (Show der Naturwunder), Pegelverhältnisse Referenzmischung	38
3	Ausschnitt (Show der Naturwunder), Pegelverhältnisse Variante 7 LU130	39
4	Reihenfolge der 16 Ausschnitte	44

1 Einleitung

„Sprachverständlichkeit“ ist ein Begriff, der im Alltag selten benutzt wird und selbst im Duden nicht zu finden ist. Trotzdem wird er in der deutschen Sprache verwendet und jeder weiß, was damit gemeint ist. Die Zusammensetzung aus den beiden Wörtern „Sprache“ und „Verständlichkeit“ ergibt Sprachverständlichkeit, also Sprache, die verständlich ist. Auch in den „Technischen Richtlinien – HDTV zur Herstellung von Fernsehproduktionen für ARD, ZDF und ORF“, die durch schriftlich festgelegte Vorgaben alle Sendungen auf ein technisch gleiches Niveau bringen sollen, wird Sprachverständlichkeit im Kapitel 16 *Ton* erwähnt:

„Für eine fernsehtaugliche Fassung muss das Mischungsverhältnis immer zugunsten der Sprachverständlichkeit gewählt werden.“¹

Im weiteren Verlauf der Richtlinien geht es um tontechnische Details einer Produktion wie das Audio-Format, Metadaten, Tonspurbelegung, Lautheitsaussteuerung oder den Versatz zwischen Bild und Ton einer Sendung. Sprachverständlichkeit stand also bisher noch nicht so sehr im Fokus, da es in diesen Richtlinien nicht weiter behandelt wird.

Inzwischen wurde jedoch eine Arbeitsgruppe der Fernsehbetriebsleiterkonferenz gegründet, die das Thema Sprachverständlichkeit im November 2012 in ihre Agenda aufgenommen hat. Grund dafür sind die zunehmenden Zuschauerbeschwerden, die sich auf die schlechte Sprachverständlichkeit in Sendungen und Formate beziehen.

So fordert Klaus Dittmann (Mitglied des Deutschen Schwerhörigenbundes) in einer Stellungnahme alle Rundfunkanstalten auf, die Sprachverständlichkeit in Sendungen für leicht- und mittelschwerhörige Zuschauer/Zuschauerinnen² zu verbessern.³ Er weist explizit darauf hin, dass es dabei nicht um schwer hörgeschädigte Personen geht. Dafür steht das Angebot des „barrierefreien Fernsehens“⁴ zur Verfügung, das immer weiter ausgebaut werden soll. In diesem Zusammenhang ist der Aspekt nicht zu vernachlässigen, dass das menschliche Gehör mit zunehmendem Alter abnimmt und bestimmte Defizite aufweisen kann. So

¹ IRT, „Technische Richtlinien – HDTV“,

<http://www.irt.de/de/publikationen/technische-richtlinien.html>, S. 24 (08.01.2014).

² Anm. d. Verf.: Im weiteren Verlauf der Arbeit wird ausschließlich eine maskuline Form für eine bessere Lesbarkeit verwendet. Die Begriffe beziehen sich grundsätzlich auf beide Geschlechter.

³ vgl. Dittmann, Klaus: <http://www.schwerhoerigen-netz.de/UMFRAGE2013> (13.01.2014).

⁴ Anm. d. Verf.: „Barrierefreies Fernsehen“ beschreibt die Verbesserung der Situation für hör- und sehbehinderte Menschen, die nur durch Untertitel oder Audiodeskription das Fernsehen ausreichend nutzen können.

nehmen ältere Menschen hohe Frequenzen schlechter wahr als tiefe Frequenzen, wodurch das Verstehen der Sprache beeinträchtigt werden kann.⁵ Bei Fernsehsendungen entstehen dann Schwierigkeiten, wenn der Sprache zusätzlich Musik und Geräusche hinzugefügt werden.

Im Jahr 2010 gab es in England ein Projekt der British Broadcast Corporation (BBC) und der beiden Organisationen Voice of the Listener and Viewer (VLV) und Action hearing loss (RNID), das einen ähnlichen Auftrag hatte wie Klaus Dittmann ihn fordert. Damals wurden ca. 20.000 Zuschauer in ganz England zum Thema Sprachverständlichkeit befragt.⁶ Das Ergebnis der Studie war, dass Menschen ab 65 Jahren große Schwierigkeiten haben, die Sprache im Fernsehen zu verstehen. In England betrifft dies fast 20 % der Bevölkerung. Außerdem wurde in dieser Untersuchung festgestellt, dass die größten Probleme fremde Dialekte oder Akzente, zu laute Hintergrundmusik und undeutlich bzw. zu schnell sprechende Schauspieler sind. Die BBC reagierte auf diese Erkenntnisse mit den Best practise tips, die im Internet frei verfügbar sind.⁷ Dort kann man Empfehlungen und Ratschläge bezüglich bestimmter Drehsituationen oder zur Hintergrundmusik nachlesen, die zu einer besseren Sprachverständlichkeit führen sollen. Ob diese Vorgehensweise die Situation im englischen Fernsehen verbessert hat, ist zu diesem Zeitpunkt noch nicht bekannt.

Da es in Deutschland eine ähnliche Problematik gibt, soll diese Arbeit in Kooperation mit dem Südwestrundfunk Stuttgart ein Ansatz zu einer *Production Guideline* sein, die zu einer besseren Sprachverständlichkeit im Fernsehen führen soll. Mit Empfehlungen und Hinweisen für Tonmeister und die Verantwortlichen des Programms sollen damit die Erwartungen der Zuschauer besser erfüllt werden. Inhalt dieser Arbeit ist hauptsächlich die Tonmischung, die für das Fernsehen eine große Rolle in Bezug auf das Verhältnis zwischen Sprache und Hintergrundmusik bzw. Hintergrundgeräusche spielt. Ein weiterer Aspekt sind die Schauspieler, die nach Aussagen der Zuschauer immer undeutlicher sprechen.

Die Grundlagen zu diesen Teilaspekten werden in Kapitel 2 behandelt, das sich mit den Themen „Sprechwissenschaften“ und „Auditive Wahrnehmung“ beschäftigt. Der Hauptteil der Arbeit besteht in einer Analyse von Zuschauerbeschwerden, die von der Arbeitsge-

⁵ vgl. Zwicker, Eberhard et al: Das Ohr als Nachrichtenempfänger, S. 55.

⁶ siehe Anhang 6.4.2, *Press Release VLV Audibility Project*.

⁷ vgl. Cohen, Danny:

<http://www.bbc.co.uk/academy/production/article/art20130702112135255> (13.01.2014).

meinschaft der öffentlich-rechtlichen Rundfunkanstalten der Bundesrepublik Deutschland (ARD)⁸ und dem Zweiten Deutschen Fernsehen (ZDF)⁹ zur Verfügung gestellt wurden, sowie einem Hörvergleich, bei dem 21 Probanden unterschiedliche Mischungsvarianten bewerten mussten. Die Schwierigkeit des Themas Sprachverständlichkeit besteht darin, dass es keine objektive Möglichkeit gibt, eine Mischung technisch auf Sprachverständlichkeit zu überprüfen. Bisher können nur über subjektive Aussagen Untersuchungen durchgeführt werden, die qualitativ ausgewertet werden.

Ziel dieser Arbeit ist es, mit Hilfe einer Analyse von Zuschauerbeschwerden, die Probleme der Sprachverständlichkeit im Fernsehen auszuarbeiten und Lösungsansätze zur Entwicklung einer *Production Guideline* anzubieten.

⁸ Anm. d. Verf.: Die ARD ist ein Zusammenschluss von neun öffentlich – rechtlichen Landesrundfunkanstalten in Deutschland.

⁹ Anm. d. Verf.: Das ZDF ist die zweitgrößte Rundfunkanstalt Deutschlands.

2 Grundlagen

In folgendem Kapitel werden die Themen „Sprechwissenschaften“ und „Auditive Wahrnehmung“ erläutert. Sie dienen als Grundlage für das in dieser Arbeit untersuchte Thema „Sprachverständlichkeit im Fernsehen“ und sind für das Verständnis des Hauptteils von Bedeutung.

2.1 Sprechwissenschaften

Die Wissenschaft des Sprechens beinhaltet alle Bereiche der mündlichen Kommunikation. Es geht dabei um die Sprechbildung, die Sprecherziehung, sowie Teilbereiche der rhetorischen, ästhetischen und therapeutischen Kommunikation.¹⁰

Wichtig für die Sprachverständlichkeit im Fernsehen sind die Kapitel ästhetische Kommunikation, Sprechbildung und Sprecherziehung. Die rhetorische und therapeutische Kommunikation tangieren andere Bereiche der Sprechwissenschaft und werden deshalb in der vorliegenden Arbeit nicht behandelt.

Die Sprache ist für den Menschen das wichtigste Kommunikationsmedium. Ohne die Möglichkeit des Dialoges, der Konversation oder Diskussion fällt es Menschen schwer Informationen oder Befindlichkeiten untereinander auszutauschen.¹¹ Die Definition des Begriffs „Sprache“ scheint einfach zu sein, doch „[...] *was meinen wir, wenn wir von Stimme, Stimmorgan, Sprechen, Singen oder Stimmschall reden?*“¹² Zur Klärung des Begriffs Sprache ist ein kurzer Überblick über die Bereiche ästhetische Kommunikation, Funktion der Stimmorgane und Sprecherziehung notwendig, um Grundlagen für die Sprachverständlichkeit zu schaffen.

2.1.1 Ästhetische Kommunikation

Ein Teilgebiet der mündlichen Kommunikation ist die ästhetische Kommunikation. In der Sprechkunst wird das Sprechen von literarischen Texten in Theorie und Praxis erarbeitet und gegebenenfalls neu interpretiert, sodass diese neben der schriftlichen / visuellen Darstellung auch auditiv wahrgenommen werden kann.¹³ Im folgenden Kapitel werden die Grundlagen der ästhetischen Kommunikation behandelt, damit neben der Sprechbildung die theore-

¹⁰ vgl. Papst-Weinschenk, Marita: Grundlagen der Sprechwissenschaften und Sprecherziehung, S. 13.

¹¹ vgl. Lazarus, Hans et al.: Akustische Grundlagen sprachlicher Kommunikation, S. 1.

¹² Sundberg, Johan: Die Wissenschaft von der Singstimme, S. 9.

¹³ vgl. Lemke, Siegrun: Sprechwissenschaft/ Sprecherziehung, S. 23.

tischen Grundlagen der Sprechwissenschaft beleuchtet werden und damit mehr auf das Kapitel „Sprecherziehung eines Schauspielers“ Bezug genommen werden kann.

Zunächst soll der Begriff „Ästhetik“ kurz erläutert werden. Henrik Nikula deutet den Begriff Ästhetik in Bezug auf literarische Texte so, dass es neben der akustischen Wahrnehmung auch um ein „*sprachliches Kunstwerk*“¹⁴ geht. Ästhetik bezeichnet demzufolge das subjektiv „Schöne“ und kann daher keinen Anspruch auf Objektivität erheben. Zum einen geht es um die „*Trennung von objektiv feststellbarer ‚artistischer Formung‘ der Sprache [...] und der subjektiven Veranstaltung einer ästhetischen Lektüre, die den Text als Kunstwerk [...] konstituiert andererseits.*“¹⁵

Das Gebiet der ästhetischen Kommunikation ist ursprünglich auf das „*Dichtungssprechen und [...] Sprechen des Schauspielers*“¹⁶ bezogen. Es geht dabei um das Erlernen von literarischen Texten, die einem Publikum möglichst ästhetisch vorgetragen werden. Hier sind die Bereiche Lyrik, Dramatik und Epik als Genre relevant, die vor allem in der früheren Sprechkunde von großer Bedeutung waren und eng mit den Fächern Literatur-, Theater- und später Medienwissenschaften verknüpft sind.¹⁷ Inzwischen wird die sprechkünstlerische Kommunikation als großes Gebiet mit folgenden Schwerpunkten angesehen:

*„Anleitung zum Dichtungssprechen innerhalb und außerhalb des Deutschunterrichts, Sprecherziehung für Theater und in der Schauspielausbildung, Sprecherziehung für Amateurtheater (früher Laienspiel), Sprecherziehung für künstlerische Aufgaben in Hörfunk und Fernsehen.“*¹⁸

Es gibt verschiedene Formen der ästhetischen Kommunikation. Das erste Teilgebiet ist das „Textsprechen“. Hier interpretiert der Sprecher einen Text subjektiv und trägt ihn dem zuhörenden Publikum vor. Dazu muss der Text zuerst auf Struktur und Gestalt analysiert werden, dann wird der Inhalt erarbeitet. Wichtig für die Sprachverständlichkeit ist zu allererst eine klare und deutliche Aussprache, sodass der Zuhörer das Gesprochene einfach versteht. Danach geht es um die Erörterung verschiedener Interpretationen, da der Text auf

¹⁴ Nikula, Henrik: Der literarische Text – eine Fiktion, S. 23.

¹⁵ Seidler, Herbert: Grundfragen einer Wissenschaft von der Sprachkunst, S. 51.

¹⁶ Gutenberg, Norbert: Einführung in Sprechwissenschaft und Sprecherziehung, S. 34.

¹⁷ vgl. Lemke, Siegrun: Sprechwissenschaft/Sprecherziehung, S. 24.

¹⁸ Gutenberg, Norbert: Einführung in Sprechwissenschaft und Sprecherziehung, S. 34.

den Sprecher selbst unterschiedlich wirken kann. Erst dann kann die konkrete Sprechfassung entwickelt werden. Diese ist hörer-orientiert und bildet somit eine Situation zwischen Sprecher und Hörer¹⁹, zu der es folgende rhetorische Frageformel gibt:

„Was will ich (wer) dadurch bei genau diesen Hörern (zu wem) erreichen (wozu), daß [sic!] ich genau diesen Text (wie) zu diesem Zeitpunkt (wann) an diesem Ort (wo) spreche, was veranlaßt [sic!] mich dazu (warum), was will ich mit diesem Text sagen (was), auf welchen Problembereich beziehe ich mich mit diesem Text (worüber) und wie muss ich unter all diesen Bedingungen den Text sprechen (wie).“²⁰

Eine andere Form der ästhetischen Kommunikation findet man im Schauspielerischen. Im Gegensatz zum reinen Textsprechen geht es dabei auch um darstellerisches Agieren, manchmal auch im Zusammenhang mit gestischem Sprechen erwähnt. Der Text kann vorgegeben sein, so dass neben dem Sprechen nur die Inszenierung geprobt werden muss oder aber der Schauspieler improvisiert und erarbeitet dadurch einen neuen Text.

Wie beim Textsprechen ist auch beim Theaterspielen die Kommunikation zwischen Schauspieler und Publikum von großer Bedeutung. Der Sprecher sagt nicht nur etwas, sondern führt den Text auch vor. Bereits Berthold Brecht bezeichnete das gestische Sprechen als „Werkzeug des Handelns“.²¹

Für vorgegebenen und improvisierten Text zeichnen sich ähnliche Fakten wie beim reinen Textsprechen ab:

„Schauspieler sprechen aus bestimmten Motivationen zu bestimmten Zwecken (mit bestimmten Zielen) an einem (Theater-) Ort zu einer (Theater-) Zeit über bestimmte Themen zu einem bestimmten Publikum Inhalte (was) im Medium von Sprache (Wörter und Sätze) und Sprechausdruck.“²²

Inzwischen ist der Begriff Text- und gestisches Sprechen nicht nur auf Lesungen oder Theater beschränkt. Durch die Entwicklung der Kommunikationsmedien, wie Hörfunk und

¹⁹ vgl. Gutenberg, Norbert: Grundlagenstudien zu Sprechwissenschaft und Sprecherziehung, S. 388f.

²⁰ ebd., S. 393.

²¹ Pabst-Weinschenk, Marita: Grundlagen der Sprechwissenschaft und Sprecherziehung, S. 192.

²² Gutenberg, Norbert: Einführung in Sprechwissenschaft und Sprecherziehung, S. 173.

Fernsehen, erweitert sich der Begriff mündlicher Kommunikation deutlich. Es gibt neben Lesungen und Theater die Medien Hörspiel, Fernsehsendungen aller Art und Kino. Für jedes dieser neuen Kommunikationsmittel ist eine ausgebildete Sprecherstimme von grundlegender Bedeutung. Schauspieler oder Sprecher müssen eine gute Sprecherziehung erfahren, bevor sie vor ein Mikrofon und/oder eine Kamera treten. Dabei ist nicht unwichtig, dass die Entwicklung der neuen elektronischen Medien die Aussprache stark beeinflusst. So war es schon in den Anfängen des Hörfunks Ende der 1920er Jahre üblich, dass professionelle Sprecher oder sogar Schriftsteller für das Vortragen von Nachrichten oder anderen Beiträgen engagiert wurden. Da die Technik bis dahin noch sehr störanfällig war und das Rauschen bei der Übertragung mitunter ein großes Problem darstellte, musste überdeutlich und überkorrekt gesprochen werden. Anders als in Film und Fernsehen kann beim Hörfunk auch keine visuelle Kompensation durch Mimik oder Körpersprache eingesetzt werden.²³ Die technischen Verbesserungen für das Aufzeichnen der Stimme erlaubten mit der Zeit das Sprechen „[...] ohne zu schreien [...]“²⁴.

Das Medium Fernsehen hatte zunächst den Anspruch an Filme, die eine theaterhafte Sprache als Grundlage hatten. Doch mit der Zeit ändert sich das Programm und es werden Live-Beiträge, Berichterstattungen, Talkshows, Interviews und Reality-Shows gesendet. Dadurch vermischen sich Textvorgaben mit spontaner Alltagskommunikation und das Fernsehen wird zunehmend zum Unterhaltungsmedium.²⁵ Schauspieler und Sprecher werden immer mehr dazu aufgefordert die Rollen noch intimer und realitätsnäher auszuführen, was sich auch im Sprachgebrauch offenbart. Die technische Weiterentwicklung der Übertragungsmedien erlaubt inzwischen zu einem gewissen Grad diese Lautschwächungen bzw. spontanes und schnelles Reden mit weniger Augenmerk auf die Artikulation²⁶, auf die in den nächsten Kapiteln näher eingegangen wird.

2.1.2 Sprechbildung

Im Folgenden wird zunächst erläutert, wie die Sprachproduktion beim Menschen funktioniert und welche Aspekte für das Hören wichtig sind. Die Arbeit befasst sich in einem späteren Kapitel genauer mit der Sprachverständlichkeit, deshalb sind folgende Erklärungen

²³ vgl.: Polenz, Peter von: Deutsche Sprachgeschichte Bd.3, S. 510.

²⁴ Hakkarainen, Heikki J.: Phonetik des Deutschen, S. 19.

²⁵ vgl. ebd., S. 516f.

²⁶ vgl. Hakkarainen, Heikki J.: Die Phonetik des Deutschen, S. 20f.

hauptsächlich auf das Sprechen und nicht auf das Singen bezogen, da es bei näherer Betrachtung Unterschiede gibt.

Die Phonation (Sprechbildung) ist in drei Bewegungsabläufe unterteilt, die man zusammengefasst als Ansatzrohr bezeichnet. Dieses wird oft mit dem Aufbau einer Zungenpfeife der Orgel verglichen und folgendermaßen dargestellt:

Die Atmung (entspricht dem Blasebalg einer Orgel), der Kehlkopf (Pfeife) und das Artikulationsorgan, gleichbedeutend bei der Orgel auch Ansatzrohr genannt.²⁷

Folgende Abbildung zeigt den zusammenhängenden Funktionskreis:

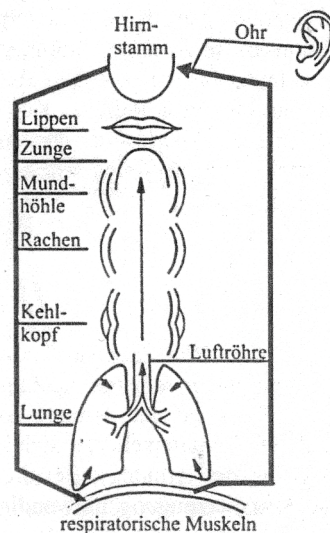


Abb. 1: Zusammenwirken der Funktionskreise.²⁸

Am wichtigsten für die Sprache und Verständlichkeit ist die Funktionsweise des Ansatzrohres. Dort werden mit dem Luftstrom und der Artikulationsmuskulatur die einzelnen Sprachlaute gebildet. Die sogenannten Artikuloren, also die beweglichen Teile des Ansatzrohres, sind: Zunge, Lippen, Unterkiefer, Gaumensegel, Rachen und Glottis (Stimmritze).²⁹ Strömt die Atemluft heraus, entsteht zunächst bei offenen Stimmlippen der sogenannte Primärklang. Dieser wird auch als „reiner Stimmtön“ bezeichnet. Obwohl er nicht wahrzunehmen ist, wird er als Schallquelle für Geräuschlaute bzw. die Sprachlaute bezeichnet. Mit Hilfe der Artikuloren wird nun der Primärklang verändert und es entstehen die Sprach-

²⁷ vgl. Habermann, Günther: Stimme und Sprache, S. 5.

²⁸ vgl. Gutenberg, Norbert: Einführung in die Sprechwissenschaft und Sprecherziehung, S. 43.

²⁹ vgl. Pompino-Marschall, Bernd: Einführung in die Phonetik, S. 43.

laute. Dazu müssen die Stimmlippen bis zu einem feinen Spalt verschlossen werden, was durch eine Unterbrechung des Luftstroms geschieht. Diese Unterbrechung führt zum sogenannten „subglottalen Überdruck“, der die Stimmlippen auseinanderdrängt und nach oben auslenkt.³⁰ Der Unterdruck, der durch das schnelle Herausströmen der Luft hervorgerufen wird, lässt die Stimmlippen wieder einwärts schwingen. Dieser Vorgang beginnt nun von vorne und bringt so die Stimmlippen zum Vibrieren bis ein Ton entsteht.³¹

Die Grafik von Johan Sundberg zeigt schematisch die Stimmschallerzeugung:

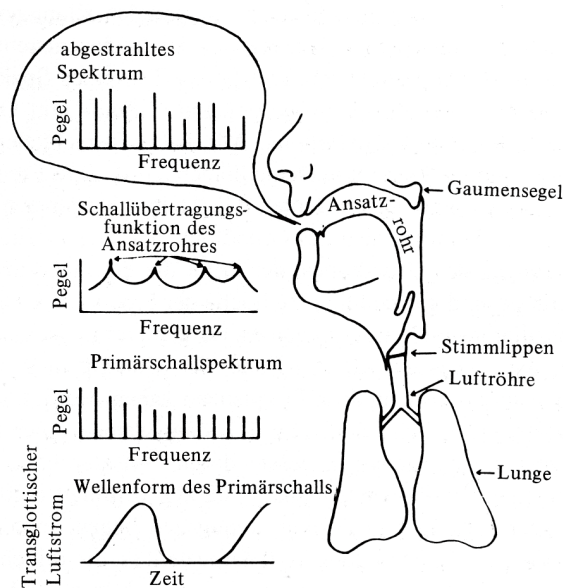


Abb. 2: Schematische Darstellung der Erzeugung von Stimmschall.³²

Die Schwingungsfrequenz der Glottis, die dabei erzeugt wird, ist die sogenannte Phonationsfrequenz (Grundfrequenz). Bei Männern und Frauen liegt diese Frequenz durchschnittlich bei 100 bzw. 200 Hz.³³ Wird der subglottale Druck wesentlich gesteigert, führt dies auch zu einer Veränderung der Grundfrequenz. Die Grundfrequenz ist der tiefste Ton, der in der Glottis entstehen kann. Durch die Resonanzräume im Ansatzrohr entwickeln sich aber viele harmonische Teiltöne, die die Phonationsfrequenz zu einem Klang und nicht zu einem einzelnen Ton werden lassen. Es gibt einige Obertöne, die durch das resonierende Ansatzrohr unterschiedlich verstärkt oder gedämpft werden. Diese Teiltöne werden Formanten

³⁰ vgl. Fiukowski, Heinz: Sprecherzieherisches Elementarbuch, S. 41.

³¹ vgl. Pabst-Weinschenk: Marita, Grundlagen der Sprechwissenschaft und Erziehung, S. 23.

³² vgl. Sundberg, Johan: Die Singstimme, S. 43.

³³ vgl. ebd., S. 61.

genannt.³⁴ Die Funktion der Formanten hat wesentlichen Einfluss auf die Sprachverständlichkeit und wird im späteren Verlauf der Arbeit unter dem Kapitel „Auditive Wahrnehmung“ näher behandelt.

Obwohl es keine „richtige“ Aussprache gibt, werden im Folgenden Kriterien einer annähernden korrekten Aussprache erläutert, die nach Siegrun Lemke den Versuch einer deutschen Standardaussprache bilden:

- *„physiologisch, somit ohne unnötige Belastungen für die Stimme,*
- *klar, somit mühelos verständlich, und zwar unabhängig von der territorialen Herkunft der Kommunikationspartner,*
- *situativ angemessen, somit variabel gebraucht,*
- *natürlich, somit weder nachlässig noch gekünstelt,*
- *ästhetisch, somit ansprechend“*³⁵

Egon Aderhold beschreibt dagegen sechs allgemeine Artikulationsmerkmale, die in der deutschen Standardaussprache berücksichtigt werden müssen:

„1. Hochrundstellung der Lippen“

Haben die Lippen eine ovale Stellung, wird der Resonanzraum vergrößert, was sich positiv auf die Klangfarbe auswirkt;

„2. Gewährleistung einer gewissen Kieferöffnungsweite“

Im deutschen gibt es keinen Laut, der mit geschlossenem Kiefer gesprochen wird; Ist der Kiefer zu geschlossen, neigt man zum Nuscheln;

„3. Zungenkontakt mit den unteren Schneidezähnen“

Bei allen Vokalen sollte die Zunge an die unteren Schneidezähne stoßen. Liegt die Zunge weiter hinten, werden die Laute dumpf;

„4. Eine gesunde nasale Setzung der Vokale und, soweit möglich, auch der anderen Sprachlauten“;

Ansonsten kann das sogenannte „Näseln“ zu einer Störung der Klangqualität beitragen;

³⁴ vgl. Papst-Weinschenk, Marita: Grundlagen der Sprechwissenschaft und Erziehung, S. 28.

³⁵ Lemke, Siegrun: Sprechwissenschaft/Sprecherziehung, S. 93.

„5. *Naturgemäße Tiefstellung des Kehlkopfes*“

Steht der Kehlkopf zu hoch, wird das Ansatzrohr verringert und führt ebenfalls zu einer schlechten Klangqualität;

„6. *Einen hygienisch, physiologischen Glottisschlag bei den Vokaleinsätzen*“

Verhindert eine gespannte und erhöhte Sprechweise ³⁶

Diese Merkmale sind eine grundlegende Voraussetzung dafür, dass neben der bewussten Artikulation keine Stimmstörungen entstehen und der Hörer von einer besseren Sprachverständlichkeit profitiert.³⁷

Das Lautsystem einer Sprache lässt sich in Vokale und Konsonanten unterteilen. Vokale oder auch Klanglaute werden „[...] *als das Fleisch, als die sinnliche Komponente der Sprache bezeichnet.*“³⁸ Sie entstehen durch das ungehinderte Ausströmen der Atemluft durch den geöffneten Mundraum. Die Artikulationsorgane bilden nun die verschiedenen Vokale „a, e, i, o und u“, sowie die Umlaute „ä, ö, ü“ und die sogenannten Diphthonge („ei, au, eu“). Im Gegensatz zu den Konsonanten entsteht bei der Bildung der Vokale kein Geräusch sondern ein Klang, weswegen die Konsonanten auch als Geräuschlaute bezeichnet werden.³⁹

Konsonanten hingegen werden durch eine Hemmstelle im Ansatzrohr artikuliert. Die Atemluft wird also an unterschiedlichen Orten im Mundraum behindert und es entsteht ein Geräusch. Bei den Konsonanten gibt es zwei Arten: zum einen „*die reinen Geräuschlaute, die völlig ohne Stimmlippenschwingung gebildet werden, zum anderen die Stimm-Geräusch-Laute, bei denen auch die Stimmlippen wie bei den Vokalen mitschwingen.*“⁴⁰ Daraus ergeben sich die stimmlosen und stimmhaften Konsonanten.

Diese Grundlagen sind Voraussetzung für eine gute Sprecherziehung. Wie Sprache gebildet wird und welche Bewegungsabläufe zu einem besseren Sprachergebnis führen, gehören zur grundlegenden Ausbildung eines jeden Schauspielers.

³⁶ Aderhold, Egon: Sprecherziehung des Schauspielers, S. 154ff.

³⁷ vgl. Lemke, Siegrun: Sprechwissenschaft/Sprecherziehung, S. 57.

³⁸ Aderhold, Egon: Sprecherziehung des Schauspielers, S. 150.

³⁹ vgl. Fiukowski, Heinz: Sprecherzieherisches Elementarbuch, S. 69.

⁴⁰ Aderhold, Egon: Sprecherziehung des Schauspielers, S. 151.

2.1.3 Sprecherziehung der Schauspieler

In der Fachliteratur werden zwei Berufsgruppen erwähnt, die eine professionelle Sprecherziehung in der Ausbildung zur Grundlage haben. Zum einen gibt es die nichtkünstlerischen, zum anderen die künstlerischen Berufe. Zu den nichtkünstlerischen Berufen zählen in der Fachliteratur z.B. alle Arten von Pädagogen (Lehrer, Erzieher, Dozenten), Politiker, sowie Prediger. Zu den künstlerischen zählen vor allem Schauspieler, wobei halbkünstlerische Berufe, wie zum Beispiel Nachrichtensprecher, sich von vorherigen unterscheiden.⁴¹ Wichtig für die vorliegende Arbeit ist die Berufsgruppe der Schauspieler, die sich bisher zum großen Teil auf den „Theaterschauspieler“ konzentrierte. Inzwischen ist der Begriff erweitert worden, und die Schauspielschulen orientieren sich auch an den neuen Medien. Besonders das Fernsehen ist in diesem Bereich extrem gewachsen, und es gibt eine Reihe von Möglichkeiten sich als Schauspieler oder Sprecher zu betätigen. Neben Fernsehfilmen gibt es Werbung, Unterhaltungsmoderation oder die Tätigkeit als Synchronsprecher. Der Begriff der ästhetischen Kommunikation erweitert also den Inhalt anspruchsvoller Literatur mit Unterhaltungs- und Informationssendungen.⁴²

Nichtsdestotrotz bleibt es für jeden Schauspieler unerlässlich, sich die Grundlagen der Sprecherziehung anzueignen, unabhängig von seinem späteren Tätigkeitsfeld.

Bisher gibt es noch nicht den einen einzigen Weg zu einer „guten“ Aussprache. Je nach Schauspieler variieren die Übungen, je nach theoretischem Vorwissen und Begabung kann an der Sprache mehr oder weniger gefeilt werden.

Vor ca. 200 Jahren war dem Schauspieler der gesamte Lebenswandel noch vorgeschrieben. Neben täglicher Gymnastik am Morgen, spazieren gehen und Stimmübungen wurde dem jungen Schauspieler geraten, sich abends nicht „in die Gesellschaft von Kollegen“⁴³ zu begeben, da Alkohol und Zigaretten der Stimme nur schaden können.

Inzwischen sind die Anforderungen natürlich nicht mehr so streng, doch ist es selbstverständlich, dass jeder Schauspieler oder Sprecher auf seinen Körper und seine Stimme achtet.

Die Sprache ist keine Einheit, die bei jedem Menschen gleich ist. Jeder Mensch hat eine andere Form der Sprachartikulation. Bei Schauspielern und Sprechern steht aber zusätzlich die Sprachverständlichkeit im Fokus. Nur durch regelmäßiges Üben und allenfalls

⁴¹ vgl. ebd., S. 15ff.

⁴² vgl. Gutenberg, Norbert: Einführung in Sprechwissenschaft und Sprecherziehung, S. 256ff.

⁴³ Aderhold, Egon: Sprecherziehung eines Schauspielers, S. 231.

auch fachliche Unterweisungen ist dieses Ziel zu erreichen. Dabei müssen nicht jeden Tag stundenlange Übungen gemacht werden. Der Experte Egon Aderhold empfiehlt entweder 30 Minuten oder 3 mal 20 Minuten am Tag gut konzentriert zu trainieren.⁴⁴

Die meisten passiven Übungen, wie das Herausziehen der Zunge oder das Drücken auf den Kehlkopf, sind inzwischen nicht mehr zeitgemäß, weswegen nur auf die aktiven Übungen eingegangen wird. Diese werden neben den atem- und körperbezogenen Techniken in sinnleere und sinnbezogene Übungen aufgeteilt.⁴⁵

Unter sinnleeren Übungen versteht man das Schulen der Stimme, die sinnbezogenen Übungen sind zuständig für eine bessere Artikulation von einzelnen Worten, Wortlauten oder kurzen sprachlichen Einheiten, die meist einen Bezug zu einem Text haben. Ganz trennen kann man beide Arten nicht, denn die Stimme wird maßgebend von der Artikulation beeinflusst.

Neben Lockerungs- und Kräftigungsübungen gibt es spezielle Übungen für die Bildung des Resonanzraumes im Ansatzrohr. Im Vordergrund steht aber das Ausbilden der Deutlichkeit und der Geläufigkeit der Stimme. Nachzusehen sind diese Übungen in der Standardliteratur für jeden Schauspieler wie z.B. „Stimmbildung auf stimm- und sprachphysiologischer Grundlage“ von Jörgen Forchhammer oder „Das ABC des Sprechens“ von Rudolf Brock.

Die beste Kontrolle für alle Übungen ist der Lehrer und natürlich das eigene Ohr. Inzwischen ist es üblich mit Aufzeichnungsgeräten die eigene Stimme aufzunehmen und durch die Hörerfahrung selbst nachzuvollziehen, woran noch gearbeitet werden muss.⁴⁶

Für den Schauspieler im Fernsehen kommt hinzu, dass er den Umgang mit einem Mikrofon beherrschen sollte. Der erfahrene Sprecher Uwe Herzog beschreibt in seinem Buch „Das Sprecherhandbuch – Ausbildung und Praxis bei Film, TV, Funk und Werbung“, dass die Stimme für das Fernsehen technisch optimal genutzt werden muss. Für ihn ist es die stimmliche Ausdruckskraft, die für die neuen Medien von Bedeutung ist.⁴⁷

Obwohl dies den meisten Sprechern und Schauspielern in Funk und Fernsehen bewusst ist, gibt es immer mehr undeutlich artikulierende Sprecher.

Die Analyse der Beschwerden in Kapitel 3 wird zeigen, dass selbst erfahrene Schauspieler zum modernen Stilmittel des „legeren Sprechens“ neigen und damit eine schlechte Sprachver-

⁴⁴ vgl. ebd., S.228f.

⁴⁵ ebd., S. 232.

⁴⁶ vgl. Aderhold, Egon: Sprecherziehung des Schauspielers, S. 242.

⁴⁷ vgl. Herzog, Uwe: Das Sprecherhandbuch, Position 412.

ständigkeit erzeugen. Die Nähe zum Publikum, die durch Umgangssprache oder schnelles Sprechen geschaffen wird, sollte nur in dem Maße stattfinden, in dem die Verständlichkeit nicht beeinträchtigt wird.⁴⁸ Die Wiener Sprecherzieherin Barbara Maria Bernhard beschreibt das Problem in der „*Angst, man könne überartikulieren.*“⁴⁹ Ihrer Meinung nach besteht der Unterschied zum Theaterschauspieler auch darin, dass beim Fernsehen kein Raum akustisch ausgefüllt werden muss. Der Schauspieler sollte feine sprecherische Mittel trainieren und über die Artikulationsspannung dem Zuschauer näher kommen. Dies ist auch notwendig, um ein Mikrofon genügend „anzuregen“. Obwohl aus akustischer Sicht kein Theaterraum gefüllt werden muss, kann ein Mikrofon nur durch eine deutliche Artikulation die Sprache sauber in elektrische Signale konvertieren. Vor allem bei Flüstern oder sehr leise gesprochenen Texten ist eine gute Aussprache unabdingbar, obwohl die technische Audio-Entwicklung wie in 2.1.1, Seite 4 dieser Arbeit angemerkt, große Fortschritte gemacht hat.

Den Grund für eine unklare Artikulation sieht Barbara Bernhard nicht in der Absicht, sprachunverständlich zu agieren, sondern eher in der Natürlichkeit und dem hohen Sprechtempo, das von vielen Schauspielern abverlangt wird.⁵⁰

2.2 Auditive Wahrnehmung

Die auditive Wahrnehmung ist ein Begriff aus dem Gebiet der Wahrnehmungspsychologie. Es geht dabei um das Hören, das beim Menschen die Funktionen der „[...] *sozialen Kommunikation, der Wahrnehmung von Umweltereignissen, bedeutungshaltigen Hörobjekten und der Wahrnehmung von Musik*“⁵¹ inne hat. Um das Sprechen als Kommunikationsmittel zu verwenden, muss das gesprochene Wort auch von den Ohren aufgenommen und verarbeitet werden. In diesem Kapitel geht es um die Hörfläche des Menschen sowie einzelne Aspekte der Sprachwahrnehmung.

⁴⁸ vgl. Lemke, Sigrun: Sprechwissenschaft/Sprecherziehung, S. 137.

⁴⁹ siehe Anhang 6.1.1 *Email Barbara Bernhard, 27.09.2013*

⁵⁰ ebd.

⁵¹ Goldstein, Bruce E.: Wahrnehmungspsychologie, S. 374.

2.2.1 Hörfläche des menschlichen Gehörs

Das Ohr als Hörorgan ist das Medium, mit dem die Sprache wahrgenommen und umgesetzt wird. Zunächst wird der Schall über die Ohren aufgenommen und im Inneren des Kopfes in elektrische Signale umgewandelt, die an unser Gehirn in Form von Impulsen weitergegeben werden.

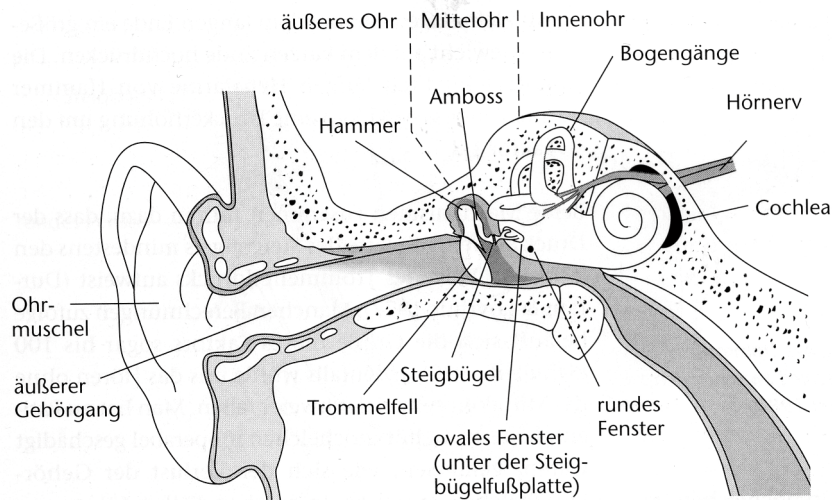


Abb. 3: Das Ohr mit seinen drei Teilen⁵²

Der Schall wird über die Ohrmuschel, den äußeren Gehörgang und das Mittelohr bis zum Innenohr in Form von Schwingungen übertragen. Die drei Gehörknöchelchen Hammer, Amboss und Steigbügel geben diese Schwingungen an das sogenannte ovale Fenster ab, das die Verbindung zum Innenohr darstellt. Das ovale Fenster ist eine Membran wie das Trommelfell und versetzt durch die abgegebenen Schwingungen die Flüssigkeit innerhalb der Cochlea in Bewegung. Diese Druckschwankungen werden im sogenannte „Cortiorgan“ in neuronale Signale umgewandelt.

Diese Signale werden vom Menschen als Hörereignis wahrgenommen, sind jedoch nicht mehr als Frequenzen oder Signale konkret messbar. Das Hörereignis wandelt sich in eine Empfindung um, die jeder Mensch für sich subjektiv bewertet und von jedem unterschiedlich wahrgenommen wird. Die Hörfläche des menschlichen Gehörs ist also nicht konkret messbar, sondern kann nur über Aussagen von Versuchsperson eingegrenzt werden. Dieser

⁵² vgl. ebd, S. 389.

Bereich ist abhängig von der Tonhöhe und dem Schalldruckpegel.⁵³ Die untere Grenzfrequenz, also der tiefste Ton, den ein Mensch wahrnehmen kann, liegt bei ca. 16 Hz, die obere Grenzfrequenz liegt bei ca. 20 kHz⁵⁴, wobei diese sehr hohen Töne zwischen ca. 17 und 20 kHz nur von einem Säugling gehört werden können.

Die Hörschwelle und Schmerzgrenze beschreiben jeweils, ab welchem Schalldruck ein Ton noch zu hören ist, bzw. ab wie viel Schalldruck ein Ton zu einer schmerzenden Empfindung übergeht. Als durchschnittliche Referenz wird der Wert von $2 \cdot 10^{-5}$ Pa bei 1000 Hz genommen. Das bedeutet also, dass bei 0 dB die Hörschwelle ansetzt, bei ca. 134 dB die Schmerzgrenze.⁵⁵

Innerhalb dieses Bereiches spielt der Zusammenhang von Frequenz und Schalldruckpegel eine große Rolle, da der Mensch bei gleichem Schalldruck aber unterschiedlicher Tonhöhe eine unterschiedliche Lautstärkenempfindung hat. Diese Beziehung wurde in den sogenannten „Kurven gleicher Lautstärke“ deutlich gemacht. An diesen kann man erkennen, wie sich das Verhältnis von Frequenz und Schalldruckpegel auf die wahrnehmende Lautstärke auswirkt. In diesem Fall wird nicht mehr von Lautstärke, sondern von Lautheit gesprochen.

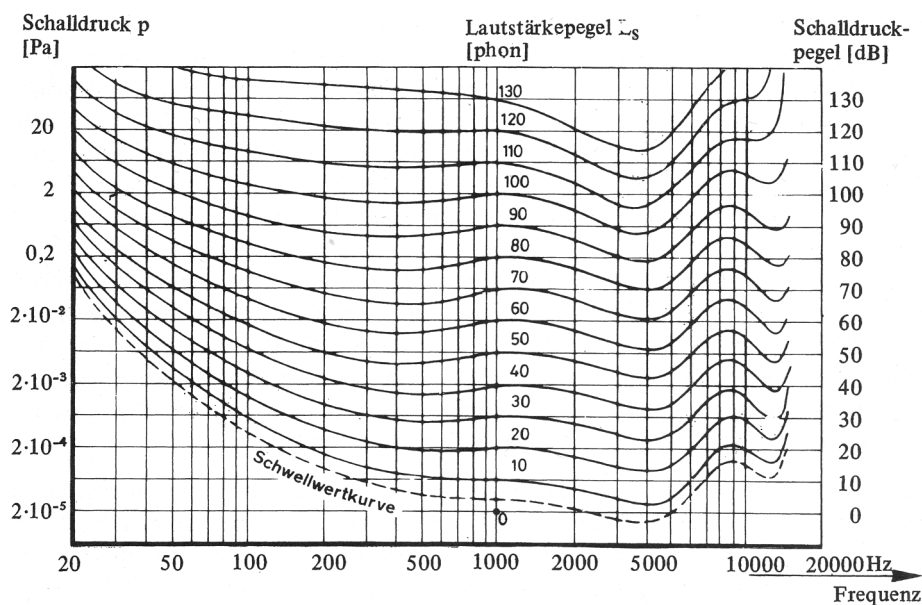


Abb. 4: Hörschwelle und Kurven gleicher Lautstärke.⁵⁶

An den Kurven ist zu erkennen, wie sich bei zunehmender Frequenz und gleichbleibendem

⁵³ vgl. Dickreiter, Michael: Handbuch der Tonstudioteknik Bd.1, S. 108ff.

⁵⁴ vgl. Flückiger, Barbara: Sounddesign, S. 199f.

⁵⁵ vgl. Görne, Thomas: Tontechnik, S. 119f.

⁵⁶ vgl. Dickreiter, Michael: Theorie der Tontechnik Bd.1, S. 111.

Schalldruck die Empfindung der Lautstärke ändert. Es zeichnet sich eine leichte Anhebung bei ca. 1000 kHz und zwischen ca. 3000 und 5000 Hz ab. Dort wird also genau der Bereich verstärkt wahrgenommen, der für die Sprachwahrnehmung am wichtigsten ist.⁵⁷

2.2.2 Frequenzspektrum der Sprache

In diesem Kapitel soll der Frequenzumfang der Sprache bestimmt werden. Sie spielt eine große Rolle in Bezug auf die Mischung der einzelnen Stems⁵⁸ einer Sendung.

Es gibt, anatomisch bedingt, Unterschiede zwischen männlichen und weiblichen Sprechstimmlagen. Abhängig von der Länge der Stimmbänder und des Vokaltrakts kann man das Spektrum beider Geschlechter von ca. 100 bis 10 kHz zusammenfassen.⁵⁹

Bei Männern liegt die Grundfrequenz der Sprache zwischen 120 und 160 Hz, bei Frauen zwischen 220 und 330 Hz.⁶⁰ Dies ist aber nur die Grundschiwingung der Stimmlippen. Durch die Vokale erweitert sich das Spektrum enorm, da manche Frequenzen durch den Resonanzraum unterschiedlich verstärkt hervorgehoben werden. Die Sprachverständlichkeit ist also nicht nur vom Spektrum der Grundschiwingung abhängig, sondern auch von den Formanten (siehe Kapitel 2.2.3 Formanten). Diese Frequenzbereiche sind die *„unerläßliche [sic!] Voraussetzung der akustischen Perzeption der Sprachlaute“*⁶¹.

Auch die Konsonanten, die grundlegend etwas höher liegen als die Vokale, tragen zu einer guten Sprachverständlichkeit bei. Bei einer Dämpfung von Frequenzen, die z.B. für die stimmhaften Konsonanten zuständig sind, kann es zu Problemen in der Unterscheidung einzelner Worte kommen. So könnte z.B. „dich“ von „sich“ oder „mich“ nicht mehr unterschieden werden.

Zusammenfassend wird in folgender Grafik dargestellt, wie sich die Vokale und Konsonanten auf das Frequenzspektrum der Sprache verteilen. Hier wird zwar deutlich, wie breit dieses Spektrum ist, trotzdem bleibt es bedeutend kleiner als das Frequenzspektrum der Musik:

⁵⁷ vgl. Flückiger, Barbara: Sounddesign, S. 200.

⁵⁸ Anm. d. Verf.: Stems sind die vorgemischten Tonspurebenen: Originalton, SFX (Geräusche + Atmos), Musik. Aus diesen Stems wird die endgültige Mischung erstellt.

⁵⁹ vgl. Lazarus, Hans et al: Akustische Grundlagen sprachlicher Kommunikation, S. 51.

⁶⁰ vgl. Dickreiter, Michael: Theorie der Tontechnik Bd.1, S. 63.

⁶¹ Habermann, Günther: Stimme und Sprache, S. 77.

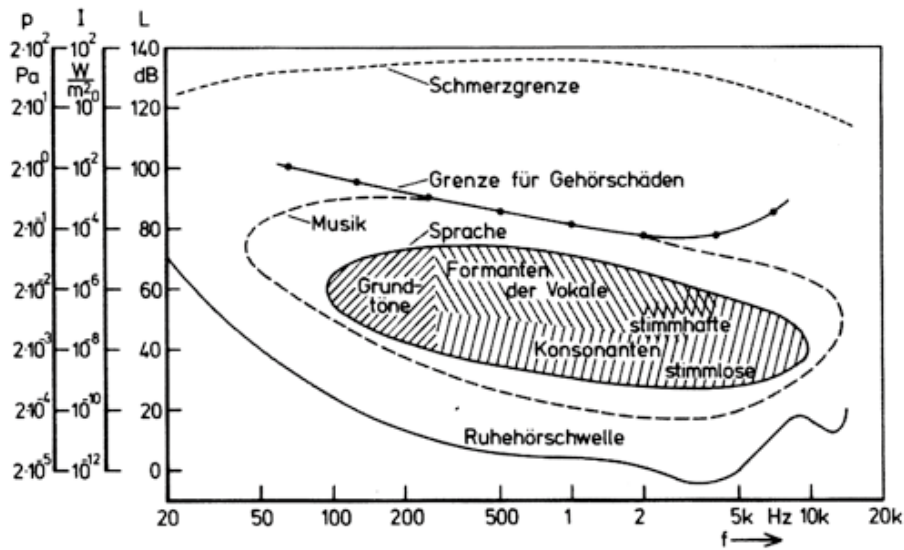


Abb. 5: Intensitäts (Pegel)- Frequenzbereich von Sprache und Musik.⁶²

Neben den Formanten spielt auch die Sprechlautstärke eine große Rolle im Bezug auf die Teiltöne, denn „je lauter gesprochen wird, um so obertonreicher ist das Sprachspektrum.“⁶³

In folgender Grafik von Manfred Spitzer wird angedeutet, wie sich das Spektrum im Vergleich zur Lautstärke ändert:

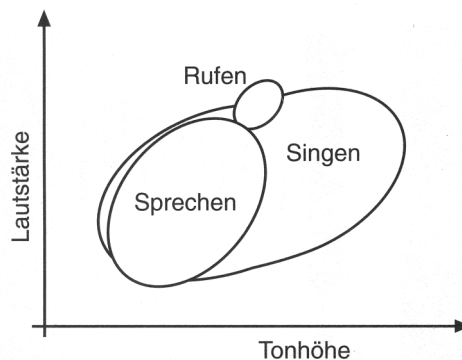


Abb. 6: Schematische Darstellung des Sprech-, Ruf- und Singstimmfeldes.⁶⁴

Die Abbildung zeigt, dass das Sprechen eine größere Dynamik als das Rufen besitzt, jedoch eine kleinere als das Singen, das auch die Tonhöhe erweitert. Das Frequenzspektrum ist also von vielen Parametern abhängig und führt deswegen zu einem stetig veränderten Frequenz-

⁶² Lazarus, Hans et al: Akustische Grundlagen sprachlicher Kommunikation, S. 19.

⁶³ Dickreiter, Michael: Theorie der Tontechnik Bd.1, S. 64.

⁶⁴ vgl. Spitzer, Manfred: Musik im Kopf, S. 256.

bereich. In folgender Abbildung wird gezeigt, wie viel Raum im Frequenzspektrum die Sprache einnimmt. Außerdem kann man schon erkennen, dass die Musik, die in Kapitel 3 eine Rolle spielt, ein sehr ähnliches Frequenzspektrum aufzeigt.

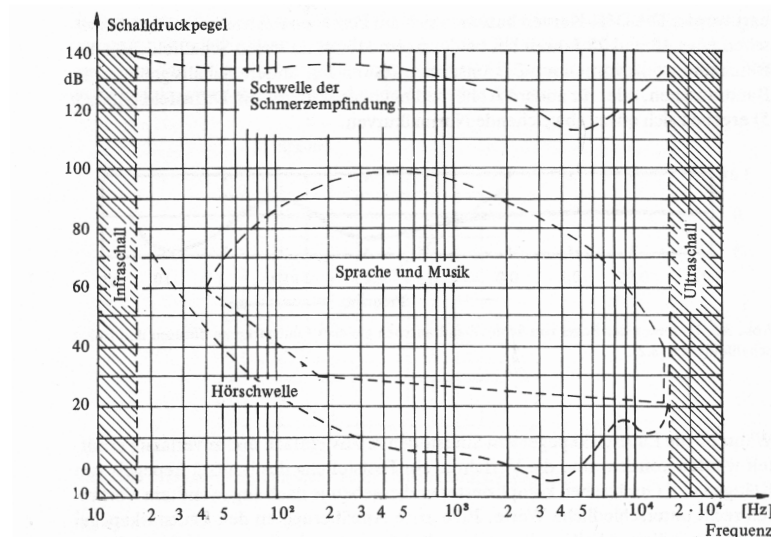


Abb. 7: Hörfeld.⁶⁵

2.2.3 Formanten

Formanten sind Frequenzbereiche, die durch Pegelanhebungen bestimmter Obertöne den Vokalen zu einer bestimmten Klangfarbe verhelfen, da das menschliche Ansatzrohr kein „verlustfrei reflektierender Körper“⁶⁶ ist und somit nicht alle möglichen Teiltöne gleichmäßig vorhanden sein können.⁶⁷

Es ist hier nicht von einzelnen Frequenzen die Rede, sondern jeder Vokal hat ein oder mehrere Spektren, die in einem gewissen Frequenzbereich liegen können. Letztendlich bestimmen die Artikuloren, wo genau sich das Maximum des jeweiligen Formanten im Spektrum befindet.

Die Töne, die in diesem Formanten-Bereich liegen, werden also durch die Gegebenheiten des Ansatzrohrs besser zum Mund geleitet als andere.⁶⁸ Schauspieler können sich diese Gegebenheit sehr zum Nutzen machen. Je besser sie artikulieren, desto besser ist die Verständlichkeit der Sprache gegenüber dem Zuhörer (siehe Kapitel 2.1.3)

⁶⁵ vgl. Dickreiter, Michael: Handbuch der Tonstudientechnik, S. 111.

⁶⁶ vgl. Pompino-Marschall, Bernd: Einführung in die Phonetik, S. 108.

⁶⁷ vgl. Habermann, Günther: Stimme und Sprache, S. 76.

⁶⁸ vgl. Spitzer, Manfred: Musik im Kopf, S. 267.

Die sogenannten Hauptformanten beziehen sich auf die Vokale „i – e – a – o – u“. Die Frequenzbereiche werden in folgender Abbildung dargestellt:

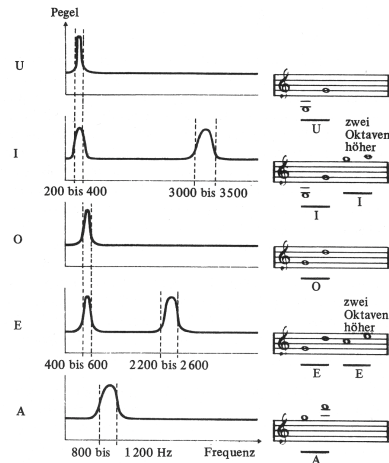


Abb. 8: Formanten.⁶⁹

An der Abbildung kann man erkennen, dass die Formanten die Grundfrequenz einer Stimme um wesentliche Frequenzbereiche erweitern. So erklärt sich auch, wie in Punkt 2.2.4 genauer erläutert, dass die Sprache durch das insgesamt breite Frequenzspektrum von anderen Hörereignissen leicht verdeckt werden kann.

2.2.4 Verdeckung

Der Begriff Verdeckung, auch Maskierungseffekt genannt (engl. masking), beschreibt die Situation des Erklings zweier unterschiedlicher Hörereignisse, infolge derer eines der Signale unhörbar wird.⁷⁰

Man unterscheidet zwischen der sogenannten Simultanmaskierung und der Vor- und Nachmaskierung.

Die Simultanmaskierung beschreibt den Vorgang, wenn ein leises Signal durch ein lauterer übertönt wird. Es erklingen zwei Töne gleichzeitig (simultan), wodurch abhängig von der Frequenz eines der beiden überdeckt wird. Das Gehirn muss dabei also zwei Signale zur gleichen Zeit aufnehmen, trennen und verarbeiten. Ist eines der Signale lauter als das andere,

⁶⁹ vgl. Dickreiter, Michael: Handbuch der Tonstudioteknik Bd.1, S. 62.

⁷⁰ vgl. Görne, Thomas: Tontechnik, S. 124.

können die Informationen des verdeckten Signals verloren gehen.⁷¹

Dabei werden grundsätzlich hohe Frequenzen von tiefen schneller maskiert, wenn die tiefen Töne einen höheren Schalldruckpegel haben als die höheren. Je näher beide Frequenzen aneinander liegen, desto größer ist der Maskierungseffekt. Je lauter der verdeckende Ton ist, desto größer ist die Bandbreite des verdeckten Frequenzbereichs.⁷²

In Bezug auf die Sprachverständlichkeit kann dies zu einem massiven Problem werden, da durch die Verdeckung die Hörschwelle nach oben verschoben wird. Man spricht dann von der sogenannten Mithörschwelle. Diese Schwelle beschreibt den Bereich, wenn ein Testsignal z.B. von Rauschen verdeckt wird und gerade noch hörbar ist.

In folgender Abbildung werden Sinustöne von weißem Rauschen verdeckt. Man kann gut erkennen, dass die Mithörschwelle nach oben wandert, je lauter das Rauschen ist. Außerdem sieht man wie bis ca. 500 Hz die Sinustöne gleichmäßig verdeckt werden, danach steigt die Mithörschwelle stetig an.

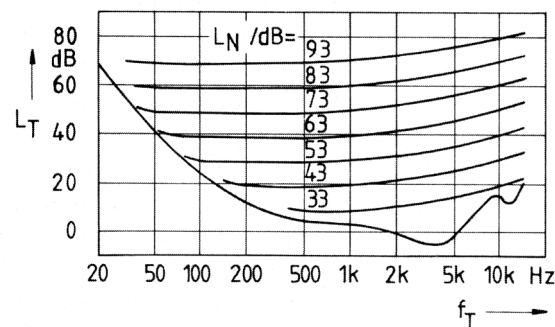


Abb. 9: Mithörschwelle.⁷³

Am empfindlichsten sind die Ohren zwischen 2 und 5 kHz. Genau dort also, wo die wichtigsten Frequenzen der Sprache liegen.

In folgender Abbildung wird ein 1 kHz Sinus - Ton mit verschiedenen Schallpegeln grafisch dargestellt. Der Ton wird jeweils um 20 dB Schritte lauter. Die Mithörschwelle steigt damit stetig an. So kann zum Beispiel bei einem Schallpegel von 80 dB des 1 kHz Tons ein 2 kHz Ton, der nur 40 dB laut ist, nicht mehr gehört werden.

⁷¹ vgl. Lazarus, Hans et al.: Akustische Grundlagen sprachlicher Kommunikation, S. 27f.

⁷² vgl. Görne, Thomas: Tontechnik, S. 124.

⁷³ vgl. Lazarus, Hans et al.: Akustische Grundlagen sprachlicher Kommunikation, S. 22.

Je größer also der Schalldruck des 1 kHz Tons, umso breiter ist der verdeckte Frequenzbereich.⁷⁴ Diesen Frequenzbereich nennt man auch kritische Bandbreite.

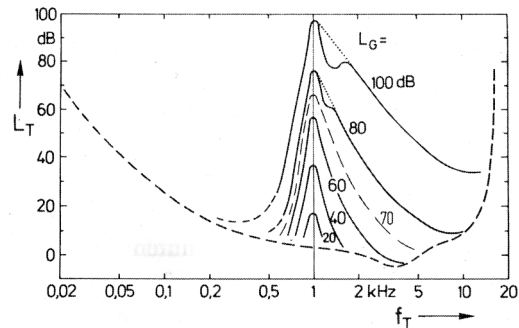


Abb. 10: Maskierung, kritische Bandbreite.⁷⁵

Speziell für die Sprachverständlichkeit gibt es den Begriff der Verständlichkeitsschwelle. Dies bedeutet, dass der Pegel der Sprache höher liegen muss, als der Pegel des Störgeräusches. Können Sprachreize gerade noch erahnt und zu ca. 50 % verstanden werden, spricht man von der Entdeckungsschwelle.⁷⁶ Beide Schwellen können in Abhängigkeit vom Pegel in folgender Grafik nachvollzogen werden.

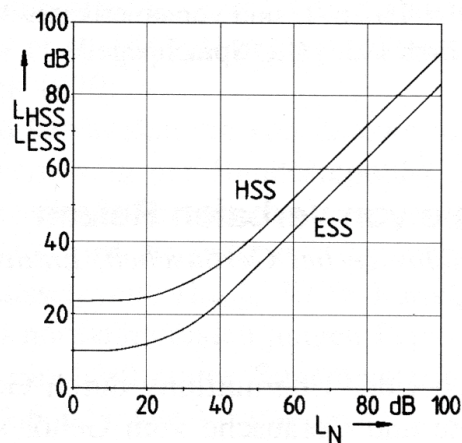


Abb. 11: Entdeckungs- und Verständlichkeitsschwelle.⁷⁷

Die Schwelle für die Verständlichkeit (HSS) liegt im Schnitt ca. 5 bis 15 dB höher als die der Entdeckungsschwelle (ESS).

⁷⁴ vgl. Görne, Thomas: Tontechnik, S. 124f.

⁷⁵ vgl. ebd., S. 125.

⁷⁶ vgl. Lazarus, Hans et al: Akustische Grundlagen sprachlicher Kommunikation, S. 28.

⁷⁷ vgl. ebd.

Der deutsche Akustiker Eberhard Zwicker beschreibt im Zusammenhang mit der Verdeckung durch Musik folgendes Beispiel: Bei einem 200 Hz Ton einer Trompete ergeben sich viele harmonische Teiltöne. An nachfolgender Grafik wird deutlich, in wieweit die ersten 10 harmonischen Teiltöne die Mithörschwelle bei 40 dB und 60 dB nach oben wandern lassen. Die Verdeckung ist also schon bei einem Ton sehr breitbandig.

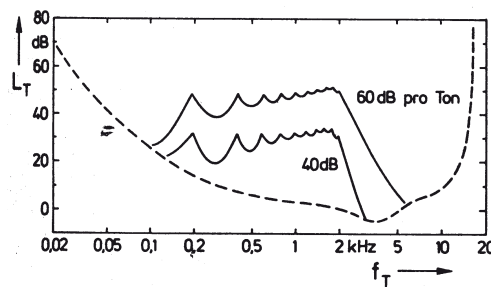


Abb. 12: Mithörschwelle verursacht durch einen 200 Hz Trompeten – Ton.⁷⁸

Ähnlich verhalten sich die Mithörschwellen bei der Verdeckung durch einen Geigenton. Wird die tiefste Saite der Geige gespielt, ergeben sich durch die Beschaffenheit der Teiltöne folgende Mithörschwellen bei 40, 60 und 80 dB Schallpegel des Violin-Tons.⁷⁹

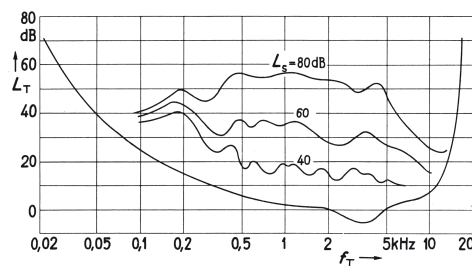


Abb. 13: Mithörschwellen eines Tones durch verschieden laut gespielte tiefe Violinklänge.⁸⁰

Die Vor- bzw. Nachmaskierung beschreibt den Vorgang, wenn sich kurz aufeinander folgende Signale stören. Vormaskierung bedeutet, dass ein Ton von einem lauterem nachfolgenden Ton, der innerhalb von 10ms eintrifft, verdeckt wird. Die Nachverdeckung funktioniert genau anders herum. Ein lauter Ton verdeckt einen maximal 20 – 30ms nachfolgenden leisen Ton.⁸¹ Diese beiden Arten von Verdeckung sind für die Thematik dieser Arbeit aber relativ unbedeutend.

⁷⁸ vgl. Zwicker, Eberhard: Psychoakustik, S. 45.

⁷⁹ vgl. Zwicker, Eberhard et al: Das Ohr als Nachrichtenempfänger, S. 65.

⁸⁰ vgl. ebd.

⁸¹ vgl. Görne, Thomas: Tontechnik, S. 125.

2.2.5 Altersschwerhörigkeit

Schwerhörigkeit ist im Allgemeinen der Begriff für das Abnehmen der Hörfähigkeit. Die häufigste Form von Schwerhörigkeit ist die sogenannte Altersschwerhörigkeit, auch sensorische Presbyakusis genannt.

Dabei handelt es sich um einen cochleoneuralen⁸² Defekt, da die Haarzellen im „Cortiorgan“ nur mehr bedingt elektrische Signale generieren können.⁸³ Die Hörfähigkeit ist je nach Alter mehr oder weniger eingeschränkt und zeichnet sich vor allem durch die Verschiebung der Hörschwelle aus.⁸⁴

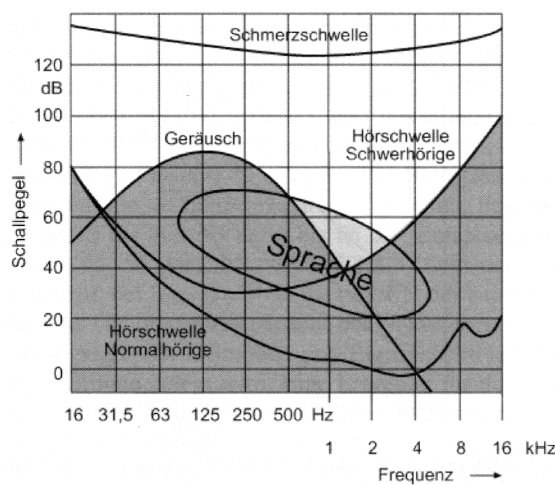


Abb. 14: Hörfläche des menschlichen Gehörs und die Einschränkung der Sprache durch Hörverlust und Geräusch.⁸⁵

Man kann deutlich erkennen, dass sich die Kurve von Schwerhörigen extrem von den Normalhörigen unterscheidet. Dabei ist der Frequenzbereich der Sprache auffallend beeinflusst. Werden zusätzlich noch bestimmte Frequenzen durch Geräusche maskiert, kann dies zu erheblichen Problemen in der Kommunikation führen.

Altersschwerhörigkeit kann schon ab einem Alter von 40 Jahren beginnen. Zunächst sind Frequenzen ab ca. 4000 Hz betroffen, die ab ca. 50 Jahren mit bis zu 30 dB Hörverlust bereits zu einer Verminderung der Sprachverständlichkeit führen können. Die Betroffenen nehmen also die Umgebung nicht nur leiser wahr, sondern können auch ein anderes Sprachempfinden haben, was sich aber nicht zwingend auf die Verständlichkeit auswirkt.

⁸² Anm. d. Verf.: cochleoneural betrifft die neuronalen Leitungen und das Innenohr.

⁸³ vgl. Goldstein, E. Bruce: Wahrnehmungspsychologie, S. 667f.

⁸⁴ vgl. Lazarus, Hans et al: Akustische Grundlagen sprachlicher Kommunikation, S. 310f.

⁸⁵ vgl. ebd, S. 311.

Die Schallpegel des Frequenzbandes der Sprache liegen zwischen 30 und 80 dB, also zum Teil weit über der Hörschwelle eines Schwerhörigen. Interessant in diesem Zusammenhang ist auch das sogenannte „Recruitment“. Es beschreibt das Phänomen, dass bei einer altersbedingten Schwerhörigkeit die Kurven gleicher Lautstärke zusammengedrängt werden. Die Schwelle, ab welchem Schallpegel etwas gehört wird, liegt zwar höher, doch die Toleranz, wann etwas zu laut ist, ist sehr klein. Dies zeigt folgende Abbildung von Gerald Fleischer:

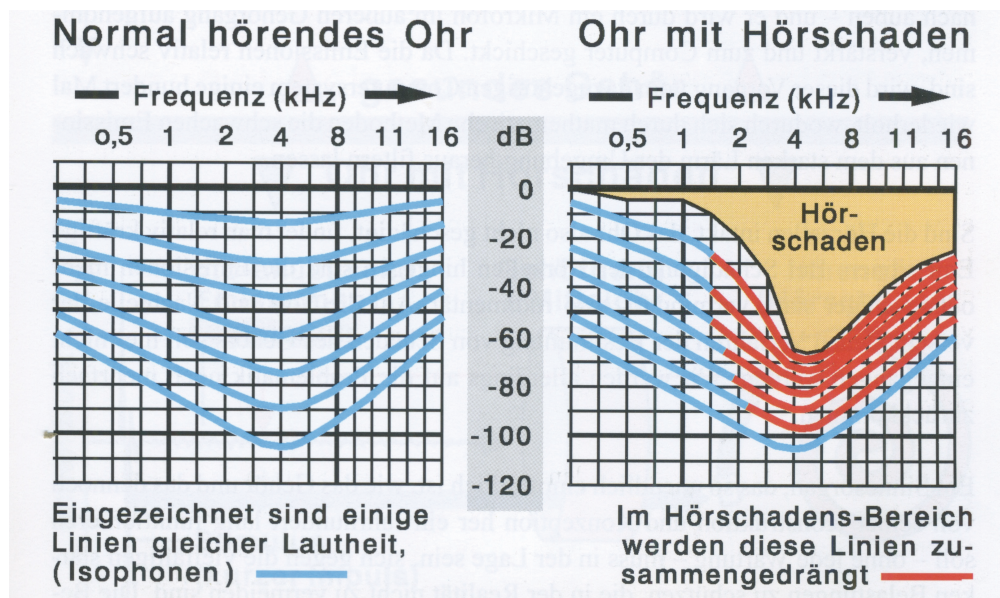


Abb. 15: Schema zum Prinzip des Recruitments.⁸⁶

Dies „erschwert die Sprachverständlichkeit außerordentlich“⁸⁷, denn: je nach Grad der Schwerhörigkeit nimmt dadurch die Fähigkeit, problemlos Sprache zu verstehen, stetig ab. Wenn sich z.B. die Hörschwelle in den Bereich der stimmhaften Konsonanten verschoben hat, wird es weitaus schwieriger einzelne Worte zu unterscheiden. Wird bei Fernsehsendungen die Sprache von zusätzlichen Hintergrundgeräuschen oder Musik gestört, hat der Hörgeschädigte große Probleme, dem Inhalt der Sendung zu folgen.

Folgende Abbildung zeigt einen durchschnittlichen Verlauf der Presbyakusis in Abhängigkeit vom Alter.

⁸⁶ Fleischer, Gerald: Gut Hören – Heute und Morgen, S. 39.

⁸⁷ ebd.

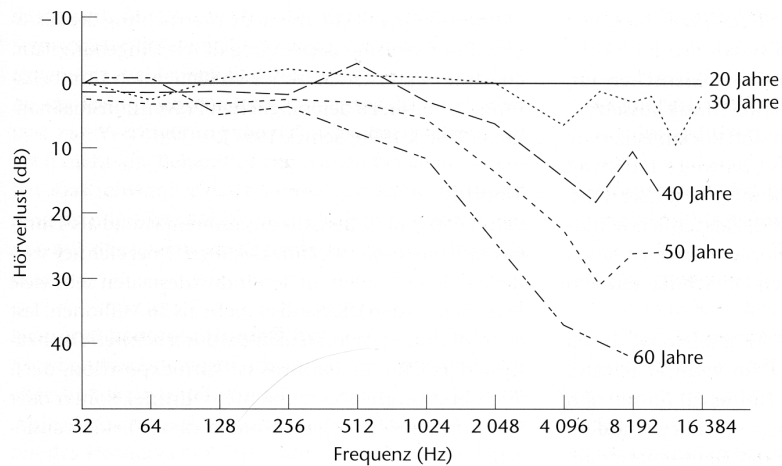


Abb. 16: altersabhängiger Hörverlust.⁸⁸

Neben dem reinen Tonaudiogramm gibt es auch ein Sprachaudiogramm, das konkret den Sprachverständlichkeitsverlust messen kann. Dieser Test ist aber enorm aufwendig, weswegen man bei Versuchsreihen über das Tonaudiogramm Rückschlüsse auf die Sprachverständlichkeit zieht.

Die in Abbildung 16 dargestellten Kurven zeigen, wie stark sich das Hörempfinden in Abhängigkeit von Alter und Frequenz ändern kann. Wird also die Kurve eines 60-jährigen betrachtet, kann der Hörverlust schon bei 1 kHz stetig sinken. Dies bedeutet nicht, dass zwangsläufig jeder 60-jährige schlecht hört. Dennoch muss man davon ausgehen, dass ein großer Teil des älteren Fernsehpublikums bis zu einem gewissen Grad an einer Hörbeeinträchtigung leidet und man sollte bedenken, dass z.B. das Durchschnittsalter des SWR – Zuschauers bei 64 Jahren liegt⁸⁹. Bei den meisten Fernsehsendungen wird überlicherweise zur Sprache noch Musik, Geräusche, Atmos⁹⁰ und Spezial-Effekte hinzugefügt, so kann man das Ausmaß des Sprachverständlichkeitsverlustes nur erahnen.

Die BBC (British Broadcast Corporation) hat zu diesem Thema bereits im Jahr 2010 zusammen mit der Voice of the Listener & Viewer (VLV), sowie dem Royal National Institute for Deaf People (RNID) eine Umfrage durchgeführt. Diese umfasste 20.000 Zuschauer, die zum Thema Sprachverständlichkeit im Fernsehen befragt wurden. Ziel war es, zu erkennen,

⁸⁸ vgl: Goldstein, Bruce E.: Wahrnehmungspsychologie, S. 667.

⁸⁹ siehe Anhang 6.4.3 *Durchschnittsalter SWR Zuschauer*

⁹⁰ Anm. d. Verf.: Barbara Flückiger erklärt den Begriff Atmos folgendermaßen: „Organisation mehrerer Klangobjekte, die einen bestimmten Schauplatz charakterisieren“. Flückiger, Barbara: Sounddesign, S. 503.

wie viele Zuschauer Großbritanniens Probleme beim Verstehen der Sprache haben und wie sich diese Probleme im Einzelnen äußern. Das Ergebnis ist eindeutig: Die Altersgruppe 65+ hat in dieser Untersuchung erhebliche Probleme, die Sprache im Fernsehen zu verstehen. Grund dafür sind zum einen die verschiedenen Dialekte, sowie die zu lauten Hintergrundgeräusche. Des Weiteren wird das undeutliche Artikulieren der Schauspieler/Sprecher, sowie das zu schnelle Sprechen genannt.

Ein großer Teil der Befragten findet die Hintergrundmusik sogar so störend, dass der Unterhaltungsfaktor des Mediums Fernsehen erheblich beeinträchtigt wird.⁹¹

Aufgrund dieser Umfrage hat die BBC die sogenannten „Best practice tips“ für die Produktion einer Sendung herausgegeben.

Schwierig ist dabei zu entscheiden, ab wann ein Zuschauer das Angebot des barrierefreien Fernsehens nutzen muss. Die Unterscheidung von hörvermindert und hörgeschädigt fällt dann schwer, wenn die Hörschädigung schon so weit fortgeschritten ist, dass ein Hörgerät benötigt wird.

In dieser Arbeit geht es ausschließlich um die Verbesserung der Situation für die Altersschwerhörigen ohne Hörhilfe, die einen großen Teil der Zuschauer der öffentlich-rechtlichen Sender ausmachen. Ungefähr eingrenzen kann man demnach das Alter von ca. 50 bis 65 Jahren. In dieser Altersgruppe kann davon ausgegangen werden, dass die meisten noch keine Hörhilfe zum Fernsehen schauen benötigen, aber dennoch leicht- bis mittelschwerhörig sein könnten.

2.2.6 Bild/Ton Beziehung

Nicht vernachlässigt werden darf die Beziehung zwischen Bild und Ton. Die Verknüpfung der beiden Sinnesreize erfolgt durch die Aufnahme zweier unterschiedlicher Informationsträger, die durch Einzel-Wahrnehmung zusammengeführt werden.⁹² Zeitlich wird das Hören schneller verarbeitet, als das Sehen. So braucht das menschliche Ohr nur ca. 3 ms, um Schall wahrzunehmen. Das Auge ist im Gegensatz dazu sehr langsam und kann optische Reize erst ab 20-30 ms im Gehirn verarbeiten. Trotzdem schafft es das Gehirn diese Laufzeit mithilfe von sogenannten „internen Taktraten“ auszugleichen, so dass der Mensch z.B. einen hörbaren Knall mit einer Waffe im Bild simultan zusammensetzt.⁹³

⁹¹ siehe Anhang 6.4.2 *Press Release-VLV Audibility Project*, S. 2.

⁹² vgl. Flückiger, Barbara: *Sounddesign*, S. 137.

⁹³ vgl. ebd., S. 139.

Der französische Komponist und Kritiker Michel Chion beschreibt in seinem Buch „Audio-Vision – Sound on screen“ das Verhältnis von Bild und Ton als sogenannte „*Synchrese*“.⁹⁴ Den Begriff setzt er aus den Worten Synchronismus und Synthese zusammen und wird von Barbara Flückiger, Sounddesignerin, als der Vorgang von „*zeitlicher Anpassung von Bild und Ton*“⁹⁵ beschrieben. Dieser Begriff drückt aus, dass beide Reize vom Menschen so verbunden werden, dass eine Einheit zwischen Bild und Ton entsteht. Diese wird vom Gehirn nicht nur als Information erkannt, sondern vom hörenden Zuschauer auch emotional verknüpft. So entstehen Gefühle und Erlebnisse, die jeder anders wahrnimmt.

In Bezug auf die Sprachwahrnehmung ergibt sich dadurch ein interessanter Effekt gerade für Menschen mit einer Hörverminderung. Durch die Verbindung von Bild und Ton entsteht ein Vorteil für Hörgeschädigte, da sie sich durch das Sehen der Lippenbewegung den Text zum Teil visuell erschließen können, wenn sie ihn nicht einwandfrei hören. Aufgrund einer Untersuchung, die in „Wahrnehmungspsychologie“ von E. Bruce Goldstein beschrieben wird, kann ein Zusammenhang zwischen Visuellem und der Sprachwahrnehmung bestehen. Zumindest wurde gemessen, dass das auditorische System durch das Lippenlesen aktiviert wurde, ohne dass letztendlich etwas gehört wurde.

Dies trifft natürlich nur auf Fernseh-Sendungen im Originalton zu. Bedenkt man die immer mehr werdenden ausländischen Filmprojekte, die synchronisiert werden, könnte dies zusätzlich zu einem Problem der Sprachverständlichkeit für ältere Menschen werden. Ein im Original amerikanischer Film mit deutscher Synchronisation kann schwieriger zu verstehen sein, da der Zuschauer die Lippenbewegung zum asynchronen Ton nicht in Verbindung bringen kann.

⁹⁴ vgl. Chion, Michel: Audio – Vision – Sound on Screen, S. 63.

⁹⁵ vgl. Flückiger, Barbara: Sounddesign, S. 141.

3 Sprachverständlichkeit im Fernsehen

Im Hauptteil dieser Arbeit werden einige Aspekte beleuchtet, die zur Verbesserung der Sprachverständlichkeit im Fernsehen beitragen können. Die im folgenden analysierten Zuschauerbeschwerden waren eine entscheidende Grundlage, um dem Problem „Mischung“⁹⁶ nachzugehen. Außerdem wurde in einem Hörversuch mit 21 Probanden eine wissenschaftliche Untersuchung zum Thema Lautstärkeverhältnis zwischen Sprache und Musik, bzw. Geräusche durchgeführt.

3.1 Zuschauerbeschwerden

In Deutschland steht für Zuschaueranfragen eine Redaktion in jedem Fernsehsenders zur Verfügung. Hier gehen Lob und Kritik der Zuschauer über das Programm und einzelne Sendungen ein. Diese Kontakte sind oft sehr aufschlussreich, da sie neben den statistischen Erkenntnissen der Gesellschaft für Konsumforschung (GfK) auch Anregungen zur Verbesserung der Programme bieten können.

3.1.1 Zuschauerredaktionen

Die jeweilige Zuschauerredaktion bekommt in der Regel täglich Anrufe, Emails oder Briefe. Diese werden nach Schlüsselworten gefiltert, bevor sie gelesen und beantwortet werden. Das Thema Sprachverständlichkeit ist eines der oft genannten Themen und wird im Ordner für „technische Anfragen“ abgelegt oder direkt an die Abteilung für Technik weitergeleitet. Dort befasst sich ein zuständiger Fachmann mit den Tonproblemen, wenn möglich sogar der zuständige Tonmeister. Für die Analyse der Beschwerden in Bezug auf die vorliegende Arbeit wurde der Kontakt zu den Zuschauerredaktionen der ARD und des SWR über die Hauptabteilung Produktion und Technik des SWR (Standort Stuttgart) hergestellt.

Über den Kontakt zu den ARD/ZDF Fernsehbetriebsleitern kamen auch Briefe, die an den WDR, NDR, MDR und das ZDF gerichtet waren.

Die ARD verschickte jede Woche das „ARD Wochenprotokoll“⁹⁷, das neben statistischen Werten auch Zitate aus Beschwerden enthält. Mit wenigen Ausnahmen wurde jede Woche mindestens eine Beschwerde zum Thema „Ton und Sprachverständlichkeit“ vorge-

⁹⁶ Anm. d. Verf.: „Mischung“ ist ein Begriff, der in der Tontechnik das Zusammenspiel von mehreren Tonspuren beschreiben. Bei Fernsehsendungen geht es darum, das richtige Lautstärkeverhältnis und den Klang von Sprache, Musik und Geräuschen in der Mischung zu finden.

⁹⁷ siehe Anhang 6.2.2 *ARD Wochenprotokolle*

legt. Dadurch kamen Informationen zusammen, die alle die gleichen Schlüsselworte wie: „undeutliche Sprache, zu laute Musik oder zu laute Geräusche“ enthielten.

Gerne wird bei Zuschauerbeschwerden das Thema „Sprachverständlichkeit“ mit dem Thema „Lautheit“⁹⁸ verwechselt. Die Werbeindustrie nutzte den bisherigen Standard der reinen Spitzenpegelmessung für Programmbeiträge aus, indem der Ton einer Werbesendung extrem komprimiert⁹⁹ wurde, aber dennoch das Pegelmaximum nicht überschritt. Das Ergebnis war, dass die Werbung wesentlich lauter als andere Sendungen wirkte, so dass zwischen den Sendungen und der Werbung erhebliche Lautstärkesprünge stattfanden. Mit Einführung der neuen EBU¹⁰⁰ Richtlinie „R128“ zum 1.9.2012 wird nun anstatt der reinen Spitzenpegelmessung die durchschnittliche Lautheit einer Sendung gemessen. Der Zielwert wird mit -23 LUFS (Loudness Units relative to full scale), bzw. 0 LU (Loudness Units, der relative Wert von LUFS) angegeben. Nach Einführung dieser Richtlinie minimierten sich die Zuschauerbeschwerden zu diesem Thema erheblich, da nun alle Sendungen in der Lautheit besser aufeinander abgestimmt sind.

3.1.2 Analyse der Beschwerden

Die insgesamt 209¹⁰¹ Beschwerden, die zum Teil schon anonymisiert waren, wurden im Zeitraum von Dezember 2011 bis Mitte November 2013 gesammelt und kategorisiert. Aufgrund der unterschiedlichen Arten der Zuschriften konnte keine lückenlose Statistik geführt werden, da die Beschwerden nicht immer alle relevanten Informationen enthielten. Dennoch konnte aus Stichworten dieser Beschwerden eine Basis für die Erstellung des Hörvergleichs (siehe 3.2) gefunden werden.

Neben dem Genre und -im besten Fall- der konkreten Nennung der Sendung, war besonders wichtig herauszubekommen, über was genau sich die Zuschauer beschwerten. Da der normale Zuschauer meist nicht zwischen Geräuschen und Atmos unterscheidet, wurden diese Bereiche zur Kategorie „Hintergrundgeräusche“ zusammengefasst. Dadurch ergab sich in der Analyse folgende Einteilung: „Hintergrundgeräusche während Sprache zu laut“ und „Musik während Sprache zu laut“. Des Weiteren haben sich viele Zuschauer über die

⁹⁸ siehe Kapitel 2.1.1, S. 14.

⁹⁹ Anm. d. Verf.: Kompression bedeutet in diesem Zusammenhang eine Dynamikeinengung und Verdichtung von Audiomaterial. Komprimierte Sprache oder Musik wirkt dadurch lauter.

¹⁰⁰ Anm. d. Verf.: Die EBU (European Broadcasting Union) ist ein Zusammenschluss von Rundfunkanstalten aus Europa, Nordafrika und Vorderasien. Diese ist für die Entwicklung und Standardisierung neuer Techniken im Bereich Fernsehen und Hörfunk zuständig.

¹⁰¹ siehe Anhang 6.2.1 A *Zuschauerbeschwerden gesamt aus Zuschriften* und 6.2.2 *ARD Wochenprotokolle*

undeutlich sprechenden Schauspieler beschwert. In diese drei Kategorien konnte nun jede einzelne Beschwerde eingeordnet werden.¹⁰²

Der „Beschwerde – Rekordhalter“ im beobachteten Zeitraum war ein Krimi, der aus der Reihe „Polizeiruf – 110“ Mitte Juli 2013 im Ersten ausgestrahlt wurde. Diese Sendung mit dem Titel „Der Tod macht Engel aus allen“ sorgte für deutschlandweites Aufsehen. Insgesamt gingen bei der ARD Zuschauerredaktion schon während der Sendung bis zum Tag danach 220 Beschwerden über die schlechte Tonqualität ein.¹⁰³ Diese Anzahl ist enorm, da in den ARD Wochenprotokollen maximal eine oder zwei Beschwerden zu einer Sendung eingehen.

Der Dialog der Schauspieler wurde von Zuschauern nicht verstanden. Besonders auf „Spiegel Online“ war die Kritik der Zuschauer enorm: „*Dramaturgie top, Sound Flop*“.¹⁰⁴

Der Bayerische Rundfunk (BR), der diese Sendung produzierte, reagierte am Tag nach der Ausstrahlung auf die vehemente Kritik und gab folgendes an:

„Man hat darauf verzichtet, den Schauspielern ihre Bewegungsabläufe kleinteilig vorzugeben.[...] Auf diese Weise konnten die Schauspieler ihren improvisatorischen Spielimpulsen spontaner folgen und wurden nicht durch enge choreografische Vorgaben gebremst.“¹⁰⁵

Der BR führte weiter aus, dass es dem Sounddepartment am Set deswegen nicht möglich war, einen technisch guten Originalton aufzunehmen. Zusätzlich zu der schlechten technischen Qualität, artikulierten die Schauspieler undeutlich, auch die Geräusche und Atmos waren gemäß der vorliegenden Beschwerden zu laut gemischt. Da diese Beschwerden im ARD Wochenprotokoll zusammengefasst mit dem Thema „schlechte Tonqualität“ notiert wurden und die 220 Beschwerden nicht einzeln vorlagen, waren sie für die Analyse zu allgemein. Sie zeigen aber dennoch das extreme Ausmaß an Beschwerden, die das Problem der schlechten Sprachverständlichkeit geäußert haben.

Es sei noch angemerkt, dass es sich bei allen Programmbreitträgen nicht ausschließlich um Eigenproduktionen des Senders handelt. Es ist auch von zugesandten Beiträgen die Rede,

¹⁰² siehe Anhang 6.2.1 B *Tabelle aller Beschwerden in Kategorien*

¹⁰³ siehe Anhang 6.2.2 *ARD Wochenprotokoll 9, S. 4.*

¹⁰⁴ kha/dpa, „Tonprobleme beim „Polizeiruf“ (2013): <http://www.spiegel.de/kultur/tv/ton-bei-polizeiruf-war-aus-dramaturgischen-gruenden-schlecht-a-911204.html> (19.09.2013)

¹⁰⁵ ebd.

wie z.B. Kinofilme, die grundlegend eine andere Dynamik aufweisen als Fernsehfilme.

3.1.3 Ergebnisse

Abgesehen von den Beschwerden bezüglich des Polizeirufs konnten insgesamt 209 Beschwerden auf die in der Analyse angegebenen Schlüsselworte eingeordnet und ausgewertet werden. Die Unterteilung in „Hintergrundgeräusche zu laut“, „Hintergrundmusik zu laut“, „Schauspieler redet undeutlich“, „Sprache nicht verstanden“ und „schlechte Tonqualität“ ergaben folgende Aufteilung in Prozent:

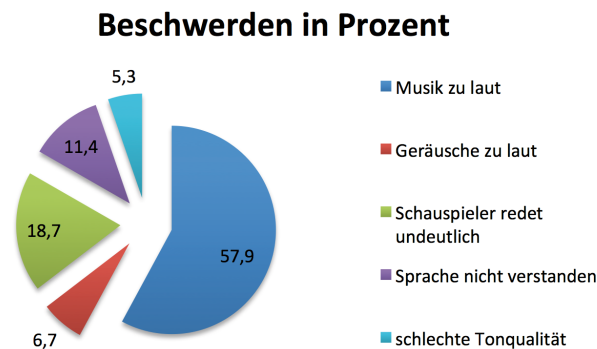


Abb. 17: Aufteilung von insgesamt 209 Zuschauerbeschwerden in Bezug auf das bemängelte Problem.¹⁰⁶

Die Tendenz ist eindeutig zu erkennen. Fast 60 % aller Beschwerden, in denen es um die Sprachverständlichkeit im Fernsehen geht, sagen aus, dass die Hintergrundmusik während der Sprache zu laut ist. Die undeutliche Sprache der Schauspieler wird um ca. drei Prozentpunkte häufiger als die lauten Hintergrundgeräusche kritisiert. Da die zu laut gemischte Musik während der Sprache überwiegt, wird dieses Problem im späteren Hörvergleich behandelt. Trotz der nicht immer genannten Sendung ist in der Kritik der Zuschauer in Bezug auf die Musik zu erkennen, dass es sich meist um die Genres „Spielfilm“ oder „Dokumentarfilm“ handelt. Dies ist einfach zu begründen, da beide Genres mit dem Medium Musik als dramaturgische Unterstützung arbeiten. Das Genre Spielfilm schließt in dieser Analyse die Genres Daily Soap und Krimi mit ein.

¹⁰⁶ siehe Anhang 6.2.3 A Beschwerden nach Problematik sortiert

Die Auswertung der Beschwerden ergab folgende Prozentanteile in Bezug auf einzelne Genres. Für die Auswertung wurden die drei wichtigsten Beschwerdegründe einbezogen:

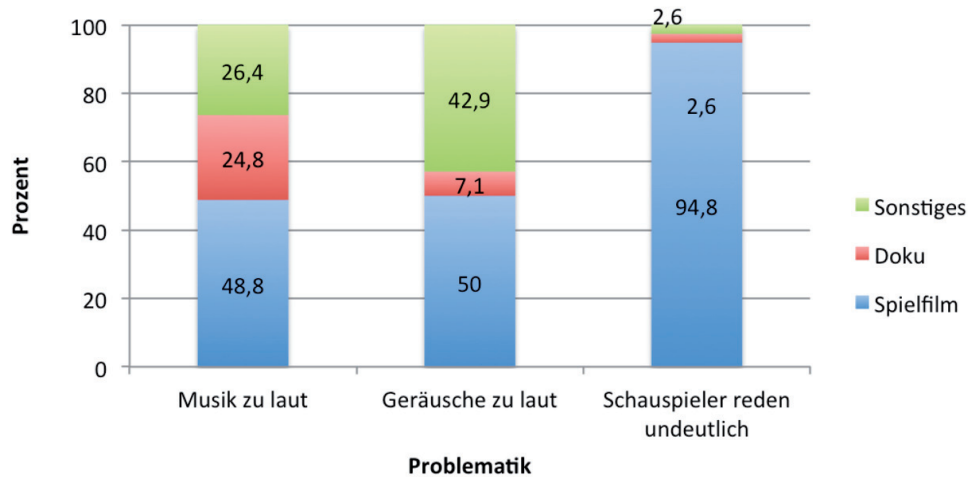


Abb. 18: Aufteilung in unterschiedliche Genres in Bezug auf die drei wichtigsten Beschwerdegründe.¹⁰⁷

Von den 121 Beschwerden bezüglich der zu lauten Musik betreffen fast die Hälfte Spielfilme, fast 25 % Dokumentarfilme bzw. Reiseberichte. Die „undeutlich sprechenden Schauspieler“ wurden zu fast 95 % bei Spielfilmen festgestellt.

Die Anzahl der Beschwerden über undeutliche Aussprache von Sprechern und Schauspielern muss näher betrachtet werden. Immer öfter werden Darsteller in Fernsehfilmen gedrängt, mehr Realität in ihrem Spiel zu zeigen, die sich nach Ansicht der Regisseure an einer „alltagssprachlichen“, eher undeutlichen Artikulation festmachen lässt. Dies bedeutet nicht, dass jeder im Alltag unklar oder verwaschen spricht, dennoch hat sich die Aussprache der Schauspieler zum Teil verändert. Mit der Entwicklung der Technik muss der Schauspieler nicht mehr überartikulieren, um verstanden zu werden.¹⁰⁸ Ein weiterer Aspekt ist auch eine deutliche Erhöhung des Sprechtempos, was bei manchen Schauspielern zu einer undeutlichen Artikulation führt.

Die knappen 17 % der Beschwerden zeigen, dass viele Zuschauer die schlampige und ihrer Meinung nach schlecht ausgebildete Aussprache der Schauspieler kritisierten, von einem

¹⁰⁷ siehe Anhang 6.2.3 B *Beschwerden nach Genre und Problematik*

¹⁰⁸ siehe Kapitel 2.1.1, S. 4.

Zuschauer auch „Til Schweiger¹⁰⁹ Syndrom“ genannt. ¹¹⁰ Laut Aussage der Redakteurin des Münchner „Polizeirufs“, ist die Improvisation der Schauspieler ein neues Stilmittel. Dies soll zu einer der Realität entsprechenden Darstellung der „Fernsehwelt“ führen. Diese Zielsetzung bringt für die Bild und Ton Aufzeichnung große Probleme: *„Die Verantwortlichen für Bild und Ton konnten die Bewegungen der Darsteller demnach nicht mehr so leicht vorhersehen und mussten spontan reagieren.“*¹¹¹

Die Beschwerden geben also eindeutig vor, was in Bezug auf Sprachverständlichkeit für das Publikum getan werden muss: bessere Aussprache, besseres Sprache-/Musikverhältnis und bessere Tonqualität.¹¹²

Interessant ist ein Vergleich der Zuschauerbeschwerden mit Aussagen der BBC Studie „Television audibility research project“¹¹³, die in Kapitel 2 unter dem Thema „Altersschwerhörigkeit“ erwähnt wurde. So stimmen die Kriterien der BBC, die eine schlechte Sprachverständlichkeit ausmachen, weitgehend mit den in dieser Arbeit überprüften Beschwerden überein. Die BBC hat in ihrer Studie die gleichen Schwierigkeiten erwähnt, die über Zuschauerbefragungen herausgefunden wurden: In dieser klagen 19 % über die fremden Akzente und Dialekte, 13 % über die zu lauten Hintergrundgeräusche (Musik und Geräusch), 14 % über die undeutliche Aussprache und 11 % über das zu schnelle Reden der Sprecher.¹¹⁴

3.2 Hörvergleich

Um für die Arbeitsgruppe Anhaltspunkte zur Entwicklung einer Production Guideline zu finden, wurde beim SWR ein Hörvergleich mit insgesamt 21 Probanden durchgeführt. Diese mussten anhand eines Fragebogens vier Filmausschnitte unterschiedlicher Genres bewerten, die in jeweils vier Mischvarianten vorgeführt wurden.

¹⁰⁹ Anm. d. Verf.: Til Schweiger ist ein deutscher Schauspieler, der neben seinen schauspielerischen Fähigkeiten auch wegen seiner undeutlichen Aussprache bekannt ist.

¹¹⁰ siehe Anhang 6.2.1 *Zuschauerbeschwerden gesamt*, S. 3 Nummer 4

¹¹¹ kha/dpa, Tonprobleme beim „Polizeiruf“, <http://www.spiegel.de/kultur/tv/ton-bei-polizeiruf-war-aus-dramaturgischen-gruenden-schlecht-a-911204.html> (19.09.2013).

¹¹² Anm. d. Verf: Konkretere Angaben zur besseren Tonqualität wurden in den Zuschauerbeschwerden nicht gemacht. Denkbar wären aber neben der undeutlichen Aussprache und der zu lauten Musik Kritikpunkte wie zu laute oder zu leise Aufnahmen, störende Hintergrundgeräusche oder auch verzerrte Wiedergabe des Tons.

¹¹³ siehe Anhang 6.4.2 *Press Release–VLV Audibility Project*

¹¹⁴ ebd.

3.2.1 Versuchsbeschreibung

Ziel des Hörvergleichs war es herauszufinden, in wie weit die Probanden Unterschiede in den Mischungen hören und ob man daraus Schlüsse ziehen kann, wie laut die Sprache gegenüber der Musik und/oder den Geräuschen sein muss. Im Fragebogen wurde deswegen nach der „Höranstrengung“ und dem „Lautstärkeverhältnis zwischen Sprache und Musik/Geräusch“ gefragt. Es wurden dafür Filmausschnitte aus vier unterschiedlichen Genres herausgesucht, die auch in den analysierten Zuschauerbeschwerden für die meiste Kritik gesorgt hatten: Krimi, Unterhaltungsshow, Sport und Dokumentation. Die Beispiele wurden von den Sendern SWR und ZDF in Form von Videodateien und vorgemischten Stems¹¹⁵ angeliefert. Im Video war auch der Stereo-Sendeton integriert, so dass eine fertige Mischung als Referenz zur Verfügung stand.

Von jedem Filmausschnitt wurden unterschiedliche Mischungen für den Hörvergleich vorbereitet. Dabei ging es immer um die verschiedenen Lautheits-Abstände zwischen Wort/Sprache zu Musik/Geräusch. Aufgrund der EBU Richtlinie „R 128“¹¹⁶ wurden alle Audio-Beispiele mit dem Lautheitsmesser TMR7¹¹⁷ gemessen. Deshalb beziehen sich alle unterschiedlichen Mischungsverhältnisse auf -23 LU bzw. 0 LUFS.¹¹⁸ entstanden insgesamt 16 Filmausschnitte, die den Probanden in gleichbleibender Zufalls-Reihenfolge vorgespielt wurden. Zu jedem Beispiel mussten die Versuchspersonen jeweils drei feststehende Fragen beantworten.

Um den Probanden ein gewohntes Bild, sowie den entsprechenden Ton zu liefern, wurden alle Beispiele über einen handelsüblichen Flachbildschirm – Fernseher¹¹⁹ dargeboten. Der Versuch fand im Hörspielstudio 13 des SWR im Funkhaus Stuttgart statt.

3.2.2 Probanden

Die insgesamt 21 Versuchspersonen wurden aus dem Umfeld des SWRs gesucht und auf folgende Altersgruppen gleichmäßig verteilt:

¹¹⁵ Anm. d. Verf.: Stems sind die vorgemischten Tonspurebenen: Originalton, SFX (Geräusche + Atmos), Musik. Aus diesen Stems wird die endgültige Mischung erstellt.

¹¹⁶ siehe Kapitel 3.1.1

¹¹⁷ Anm. d. Verf.: TMR7 ist ein Lautheitsmesser des Herstellers „Radio-technische Werkstätte“ (RTW).

¹¹⁸ Anm. d. Verf.: LU ist die Maßeinheit für Lautheit und heißt Loudness Units, LUFS bedeutet Loudness Units relative to full scale.

¹¹⁹ Anm. d. Verf.: Modell: Sony KDL-42W656A, 42”

Jahre	14 – 29	30 – 39	40 – 49	50 – 59	60 - 69	insgesamt
Anzahl	5	4	4	4	4	21

Tab. 1: Altersgruppen für den Hörvergleich.

Die Probanden wurden für den Versuch anonymisiert, so dass jede Person nur noch einer Altersgruppe zugeordnet werden musste. Alle Personen unterzogen sich außerdem einem Hörtest in Form eines sogenannten Reintonaudiogramms¹²⁰, welches das subjektive Hörvermögen jedes Probanden feststellen sollte. Dadurch wurden schwerwiegende Hörbeeinträchtigungen ausgeschlossen, die das Ergebnis des Hörvergleichs eventuell beeinträchtigt hätten. Die Probanden wurden so ausgewählt, dass sie keinen professionellen Bezug zum Medium Ton hatten. Vom Betriebsarzt des SWR, Dr. Walker, wurden die Ergebnisse des Hörtests geprüft und bei neun Probanden eine leichtgradige Schwerhörigkeit festgestellt. Drei Testpersonen wiesen eine mittelgradige Schwerhörigkeit auf und lassen sich auch in die Altersgruppe 50 plus einordnen. Die restlichen Probanden haben ein normales Hörvermögen.¹²¹ Herr Walker erklärte nochmals, dass die Auswertung der audiometrischen Tests nur bedingt etwas über das Sprachverstehen aussagen. Dazu wäre ein zeitaufwendiger Sprachverständlichkeitstest nötig gewesen, der in dieser Arbeit leider nicht möglich war. Dennoch war auszuschließen, dass einer der Probanden eine schwerwiegende Hörbeeinträchtigung aufwies.

3.2.3 Filmausschnitte

Die vier Filmausschnitte waren aus drei Produktionen des SWR, sowie einer Produktion des ZDF. Alle Beispiele waren ca. eine Minute lang und hatten jeweils Musik und Atmo unter der Sprache. Es handelte sich um Ausschnitte aus folgenden Sendungen:

„**Tatort – Spiel auf Zeit**“ (ausgestrahlt im Mai 2013, SWR)

„**Die große Show der Naturwunder – Folge 18**“ (ausgestrahlt im Juni 2012, SWR)

„**Terra X – Besonders die Katzen sind schuld**“ (ausgestrahlt im März 2013, ZDF)

„**Fussball - Länderspiel: Deutschland gegen Brasilien**“ (live Übertragung im Juni 2011, SWR)¹²²

¹²⁰ Anm. d. Verf.: Der Hörtest wurde von Frau Bosch, Arzthelferin von Dr. Walker (Betriebsarzt des SWR Stuttgart) durchgeführt

¹²¹ siehe Anhang 6.1.5 *Email Dr. Walker, 26.11.2013*

¹²² siehe Anhang 6.3.1 *Referenzvideos 6.3.1.2-6.3.1.5*

Die Szene des Tatorts wurde auf einer großen Baustelle gedreht, an der die lauten Umgebungsgereusche schon zu einer schlechten Sprachverständlichkeit führen. In der großen Show der Naturwunder erklärt der Physiker und Moderator Ranga Yogeshwar ein Experiment mit Luft, während er dieses gleichzeitig vorführt. Der Ausschnitt aus der Dokumentations – Reihe „Terra – X“ ist mit vielen Effekten unterlegt, die die Schnitte im Bild durch den Ton zusätzlich unterstützen. Der Sport-Ausschnitt stammt aus dem Länderspiel Deutschland gegen Brasilien. Zum Kommentator wird in diesem Fall die Atmo des Stadions gemischt. Die Reihenfolge der 16 Ausschnitte wurde mit Hilfe eines Randomizers generiert, damit eine zufallsorientierte Abfolge festgelegt wurde. So sollte vermieden werden, dass zwei Mischvarianten des gleichen Ausschnitts hintereinander abgespielt werden. Die Testpersonen hatten somit keine Vergleichsmöglichkeit innerhalb der vier Varianten eines Ausschnitts.¹²³

3.2.4 Herstellung der Mischvarianten

Alle Filmausschnitte wurden nun mit vier unterschiedlichen Mischvarianten unterlegt. Dazu wurde von jedem Filmbeispiel zunächst die vorhandene Stereomischung des Sendetons auf Lautheit überprüft. Gearbeitet wurde mit der Software ProTools HD 9¹²⁴ und dem Lautheitsmesser TMR7¹²⁵.

Da die Ausschnitte in Bezug auf den ganzen Film sehr kurz waren, wurden alle Beispiele außer der Fußballsending auf -23 LU korrigiert. So wurde eine Referenz geschaffen, damit alle Ausschnitte die gleiche Lautheit hatten. Anschließend wurde die Sendemischung mit den vorhandenen Stems¹²⁶ nachgemischt. Da die Originalfassungen zum Teil sehr dynamisch gemischt wurden, konnten die Lautstärkeverhältnisse zwischen Sprache und Musik/Geräusch nicht exakt nachvollzogen werden. Die Abweichungen zwischen Sendemischung und Nachmischung können aber vernachlässigt werden, da es sich nur um sehr geringe Unterschiede handelt. Durch das Anpassen der einzelnen Stems konnten nun die Pegel durch die Volume Anzeige in Pro Tools auf jeder Spur in dB bestimmt werden. Diese dB-Werte wurden zur Erstellung der vier unterschiedlichen Mischvarianten benutzt.

¹²³ siehe Anhang 6.4.1 *Randomizer*

¹²⁴ Anm. d. Verf.: ProTools ist eine Software der Firma „Avid“, die eine professionelle Musik- und Audiobearbeitung zulässt.

¹²⁵ Anm. d. Verf.: TMR7 ist ein Lautheitsmesser des Herstellers „Radio-technische Werkstätte“ (RTW).

¹²⁶ Anm. d. Verf.: Stems sind die vorgemischten Tonspurebenen: Originalton, SFX (Geräusche + Atmos), Musik. Aus diesen Stems wird die endgültige Mischung erstellt.

Nach dem Anpassen der einzelnen Stems wurde jeweils deren Lautheit gemessen. So entstanden beim Beitrag aus der Show der Naturwunder folgende Werte:

	Gesamter Mix	Werte in dB/ Protocols	Werte LUFS der Stems	Werte IT (Atmo + Musik)	Abstand Wort zu IT
Referenzmischung	-23,6 LUFS	Wort: -2,0 Atmo: -3,0 Musik: -13,1	Wort: -23,0 Atmo: -29,5 Musik: -31,6	-29,1	6 LU

Tab. 2: Ausschnitt Show der Naturwunder, Pegelverhältnisse Referenzmischung.

Da bei der Sprachverständlichkeit das Thema Verdeckung ein wichtiger Bestandteil ist, war es naheliegend, die Sprache bei jedem neuen Mischungsverhältnis auf dem gleichen Pegel zu lassen. So hatte man durchgehend das Wort als Referenz, das immer die gleiche Lautstärke aufwies. Nun wurden die Musik und/oder die Geräusche verändert, um so neue Mischungsverhältnisse zu generieren. Diese unterschieden sich in der integrierten Lautheitsmessung minimal zur Lautheit des ursprünglichen Mixes, da das Wort als Referenz immer sehr nahe an -23 LU war.

Zwei der vier Mischungsvarianten wurden in einem Mischlabor beim NDR¹²⁷ zusammen mit Diplom-Tonmeister Askan Siegfried erarbeitet. Als Ergebnis dieser Versuche empfahl er eine sogenannte „comfort zone“ für den Abstand zwischen Sprache und Musik/ Geräusch einzuführen. Das bedeutet, dass es eine obere und untere Grenze des Lautheitsabstandes zwischen Musik und Sprache geben könnte. Zwischen diesen Grenzen hätte der Mischtonmeister somit einen Spielraum, die Lautheit von Musik und Sprache aneinander anzugleichen, ohne dass die Musik zu leise unter dem Wort ist. Für diesen Bereich schlägt Askan Siegfried nun folgende Werte vor: Die Musik und Atmo sollte mindestens 7 LU und maximal 10 LU leiser als das Wort sein. So stünde ein Rahmen von 3 LU zu Verfügung, in dem die Musik und die Atmo dazugemischt werden können und wenig Verdeckungseffekte entstehen.¹²⁸

¹²⁷ Anm. d. Verf.: Norddeutscher Rundfunk

¹²⁸ siehe Anhang 6.1.2 *Email Askan Siegfried, 05.09.2013*

Aufgrund dieser Idee wurden zwei der vier Mischungsvarianten so festgelegt, dass bei gleich laut bleibendem Wort die IT-Spur¹²⁹ jeweils 7 bzw. 10 LU leiser als die Sprache war. Dies wurde durch gleichmäßiges Verändern der Pegel von Musik und Geräusche erzielt und durch Messung der Lautheit der IT Spur auf den gewünschten Wert gebracht.

Am Beispiel „Die große Show der Naturwunder“ veränderten sich die Pegel folgendermaßen:

	Gesamter Mix	Werte in dB/ Protocols	Werte LUFS der Stems	Werte IT (Atmo + Musik)	Abstand Wort zu IT
Variante 7 LU	- 22,4 LUFS	Wort: - 2,0 Atmo: - 4,0 Musik: - 14,0	Wort: -23,0 Atmo: -30,0 Musik: -32,5	-30,2	7 LU

Tab. 3: Ausschnitt Show der Naturwunder, Pegelverhältnisse Variante 7 LU.¹³⁰

Auf gleiche Weise wurde auch die Variante mit 10 LU unter dem Wort hergestellt.¹³¹ Die beiden anderen Mischvarianten entstanden aus der Überlegung heraus, dass es auch ein unangenehmes Mischungsverhältnis bei zu lauter Musik gibt. Für den Hörvergleich wurde der Abstand von nur 2 LU gewählt, um einen eindeutigeren Unterschied zu hören. Auf Anregung von Dr. Peter Mechtler sollte auch eine Mischung erstellt werden, bei der die „Lautstärkeanpassung“ über die Absenkung von Frequenzbändern geschehen sollte. Als Audio-Vorlage wurde die Mischvariante „laut“ verwendet: 2LU Abstand zwischen Sprache und Musik. Die Frequenzen wurden mit folgendem Verfahren ermittelt: Überprüft man nur das Wort in einem Spektrogramm, kann man die Frequenzen erkennen, die in der jeweiligen Stimme am stärksten hervortreten und dadurch anfällig für Verdeckungen sind. Diese Frequenzen kann man mit einem Equalizer in der Musik und den Geräuschen absenken. So ist der Verdeckungseffekt weniger gravierend und man kann die Musik und Geräusche trotzdem auf einem gewissen Pegel belassen.

Für die Analyse der Frequenzen wurde die Restaurations-Suite „RX 2“ von „iZotope“ verwendet. Die Funktion „Spectrum Analyzer“ kann durch das Öffnen der Datei, die analy-

¹²⁹ Anm. d. Verf.: IT bedeutet in Bezug auf Film die internationale Fassung einer Tonspur. Originaltöne, Musik und Atmo werden als „IT Band“ bezeichnet. In dieser Arbeit wird die IT Fassung als Zusammenfassung von Musik und Atmo bezeichnet.

¹³⁰ siehe Anhang 6.3.4 *Werte Mischungsvarianten*, S. 2.

¹³¹ ebd.

siert werden soll, benutzt werden. Am Beispiel „Die große Show der Naturwunder“ zeigt sich folgendes spektrale Abbild der Frequenzen, wenn ausschließlich das Wort analysiert wird. Von links nach rechts können hier die gesprochenen Worte im Zeitverlauf, von unten nach oben werden die Frequenzen von 0 bis 20000 Hertz angezeigt. Die gelben bis orange-farbenen Stellen zeigen die Intensität der vorhandenen Frequenzen: je gelber, umso lauter ist die angezeigte Frequenz in einem Wort vorhanden. Ist ein Bereich schwarz, hat zu dieser Zeit der Analyzer keine Frequenz erkannt, es wurde also nicht gesprochen.

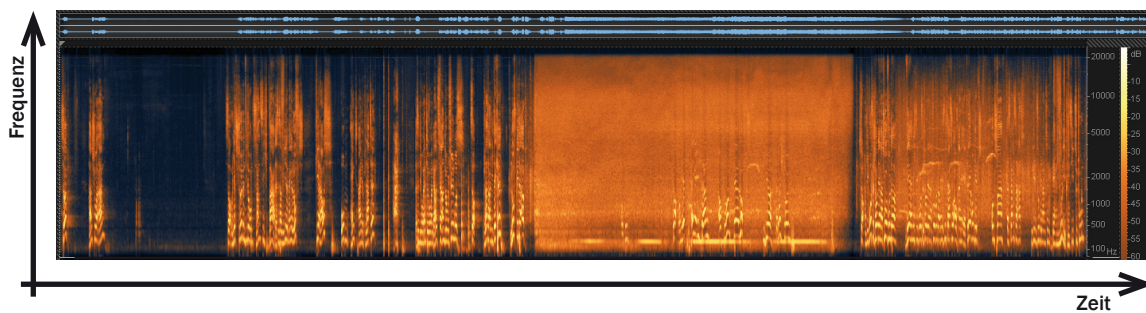


Abb. 19: Izotope Show der Naturwunder.¹³²

Man erkennt deutlich die helleren Frequenzbereiche, die auf die starken Frequenzen der Sprache hindeuten. In diesem Fall wurden die Frequenzen 450 und 1000 Hz für den Equalizer erkannt. Dies zeigt folgende hochauflösende Darstellung in iZotope:

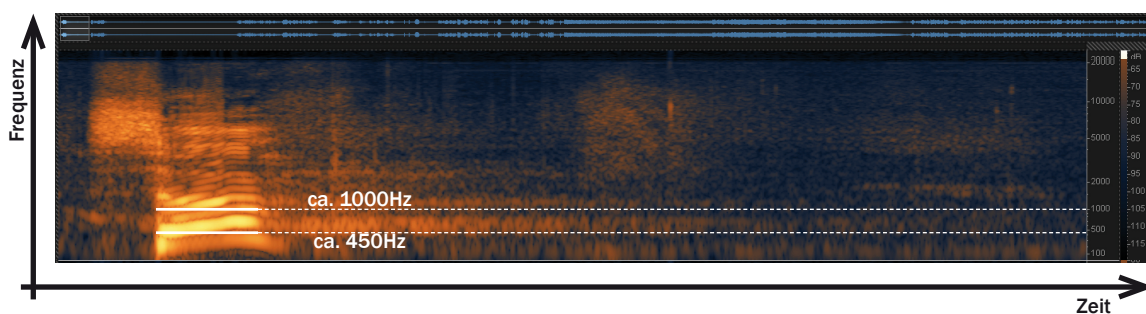


Abb. 20: Izotope Frequenzen für Equalizer eines „Show der Naturwunder“ Ausschnitts.¹³³

Für diese Version wurde der Equalizer im ProTools¹³⁴ auf diese zwei Frequenzen eingestellt. Die Güte betrug immer 1, abgesenkt wurde jeweils um 6 dB.¹³⁵ Dadurch änderte sich die

¹³² siehe Anhang 6.3.2.1 *Izotope Wort Experiment*

¹³³ siehe Anhang 6.3.2.2 *Izotope Wort Experiment EQ*

¹³⁴ Anm. d. Verf.: Der Equalizer EQ III ist in jeder ProTools Version vorhanden.

¹³⁵ Anm. d. Verf.: Diese Anregung kam ebenfalls von Dr. Peter Mechtler.

darauflin gemessene Lautheit im Vergleich zu der Fassung ohne Equalizer um fast 2 LU. Zusätzlich kann an Abbildung 19 gesehen werden, wie die im Experiment ausströmende Luft den gesamten Frequenzbereich von ca. 20 bis 20000 Hz einnimmt. So ist die Sprache rein visuell schon durch die Verdeckung schlechter zu erkennen, was sich natürlich auch auf das akustische Verständnis auswirkt.

Auf diese Weise entstanden von jedem Filmausschnitt die Fassungen:

1. IT¹³⁶ = 7 LU leiser als das gesprochene Wort
2. IT = 10 LU leiser als das gesprochene Wort
3. IT = 2 LU leiser als das gesprochene Wort
4. IT = 2 LU leiser als das gesprochene Wort + Equalizer

So waren insgesamt 16 Filmausschnitte mit jeweils vier mal vier Mischungen für den Hörversuch vorhanden.¹³⁷

3.2.5. Fragebogen

Für die Bewertung der Filmausschnitte wurde ein zweiteiliger Fragebogen erstellt. Die ersten sechs Fragen enthielten allgemeine Themen wie Alter, Geschlecht, Fernsehgewohnheiten bezüglich der Geräteausstattung zu Hause, sowie Fragen zum Problem der Sprachverständlichkeit im Fernsehen. Diese Fragen mussten nur am Anfang des Versuchs von jedem Probanden beantwortet werden. Der zweite Teil des Fragebogens bestand aus drei Fragen, die jeweils zu jedem Film-Ausschnitt beantwortet werden mussten.¹³⁸ In diesen Fragen ging es um die Höranstrengung und das Lautstärkeverhältnis zwischen Sprache und Musik bzw. Atmos.

¹³⁶ Anm. d. Verf.: IT bedeutet in Bezug auf Film die internationale Fassung einer Tonspur. Originaltöne, Musik und Atmo werden als „IT Band“ bezeichnet. In dieser Arbeit wird die IT Fassung als Zusammenfassung von Musik und Atmo bezeichnet.

¹³⁷ siehe Anhang 6.3.5 *Filmausschnitte*

¹³⁸ siehe Anhang 6.3.3 *Fragebogen*

3.2.6 Studio „13“ des SWR

Der Versuch fand im „Studio 13“ des SWR Funkhauses in Stuttgart statt. Das Studio wird für Hörspiele und Sprecheraufnahmen verwendet. Der Raum zeigt folgenden Grundriss auf:

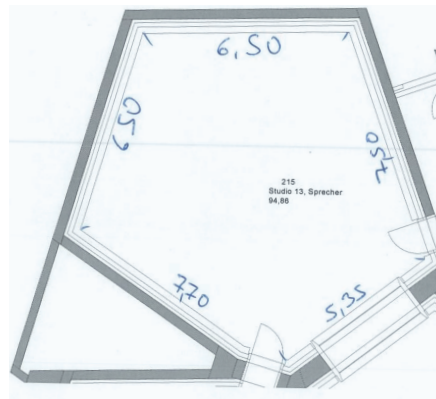


Abb. 21: Grundriss Studio 13, SWR Stuttgart.¹³⁹

Für den Versuch wurde ein Heim-Flachbild- Fernseher der Firma Sony aufgebaut, der auch als Audiowiedergabe-Gerät für die folgenden Messungen benutzt wurde.

In der Vorbereitung wurde der Raum mit dem „Acoustilyzer AL 1“¹⁴⁰ der Firma NTI auf Nachhallzeit (RT60)¹⁴¹ und Frequenzgang mit dem dazugehörigen Messmikrofon „Mini SPL“ überprüft. Mit Hilfe eines Neutrik A1¹⁴², der den Sweep für die Ermittlung des Frequenzganges bereitstellte, wurde folgender Frequenzgang für den Bereich von 375 Hz bis 17,3 kHz gemessen:

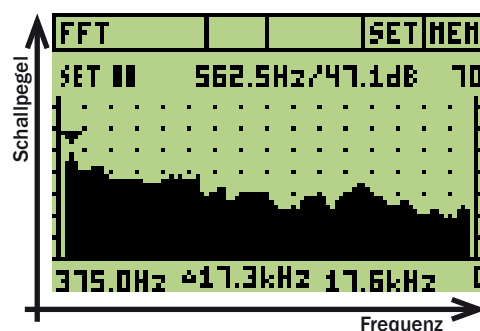


Abb. 22: Frequenzgang Studio 13 gemessen mit dem „Acoustilyzer AL1“.¹⁴³

¹³⁹ siehe Anhang 6.1.6 Email Erich Ebert, 05.02.2014

¹⁴⁰ Anm. d. Verf.: Der Acoustilyzer ist ein Audio- und Akustik Analysator.

¹⁴¹ Anm. d. Verf.: RT60 ist die Bezeichnung für das Zeitintervall, in dem der Schallpegel der abgeschalteten Schallquelle um 60dB abgenommen hat.

¹⁴² Anm. d. Verf.: Der Neutrik A1 ist ein portabler Audiomessplatz, mit dem die meisten für den Audiobereich relevanten Messungen durchgeführt werden können (z.B. Verzerrung, Pegel, FFT).

¹⁴³ siehe Anhang 6.4.4 Messungen Studio 13, S. 2.

Für die Messung der Nachhallzeit wurde der Raum mehrere Male mit rosa Rauschen über die Lautsprecher¹⁴⁴ des Fernsehers angeregt und nach kurzer Zeit wieder abgeschaltet. Diese Abschaltung erkennt der AL1 und kann nach mehreren Vorgängen die Nachhallzeit berechnen. Nach 10 Durchgängen ergab die Messung folgendes Ergebnis:

RT60	EEP	SET	MEM
STAT: STOP			
AURG	%	▶	
AURG: 10 #			
4kHz	0.27s	████	6.0%
2kHz	0.30s	████	8.0%
1kHz	0.28s	████	12.3%
500Hz	0.33s	████	16.2%
250Hz	0.31s	████	23.4%
125Hz	-.--s	---	--.-%

Abb. 23: RT60 Studio 13 gemessen mit dem „Acoustilyzer AL1“.¹⁴⁵

Der Raum zeichnet sich durch eine kurze Nachhallzeit aus, die für Sprecherstudios nach den Richtwerten von Kuhl¹⁴⁶ für eine optimale Sprachverständlichkeit sorgen.

3.2.7. Versuchsablauf

Für den Hörversuch wurden die Probanden auf fünf Termine so verteilt, dass jeweils vier bzw. fünf Personen pro Sitzung anwesend waren und jeder gut vor dem Fernseher Platz hatte.



Abb. 24: Foto eigen, Probanden in Studio 13 vor dem Fernseher.

¹⁴⁴ Anm. d. Verf.: Die Firma Sony stellte leider keinen Frequenzgang der Lautsprecher zur Verfügung.

¹⁴⁵ siehe Anhang 6.4.4 Messungen Studio 13, S. 1.

¹⁴⁶ vgl. Dickreiter, Michael: Handbuch der Tonstudioteknik Bd.1, S. 35.

Zum Abhören dienten die eingebauten Lautsprecher des Fernsehers, die wie bei den meisten Flachbildfernsehern üblich, nach hinten abstrahlen. Die einzelnen Filmausschnitte wurden über einen MacMini Server¹⁴⁷ und dem VLC – Player¹⁴⁸ abgespielt. Die Reihenfolge der 16 Beispiele wurde zuvor durch einen Randomizer¹⁴⁹ festgelegt und blieb bei jeder der fünf Probandengruppen gleich:

1	Tatort	Laut EQ	9	Experiment	7 LU
2	Experiment	Laut	10	Katzen	10 LU
3	Katzen	Laut	11	Tatort	10 LU
4	Tatort	7 LU	12	Fußball	7 LU
5	Fußball	Laut EQ	13	Experiment	10 LU
6	Katzen	Laut EQ	14	Katzen	7 LU
7	Fußball	Laut	15	Experiment	Laut EQ
8	Tatort	Laut	16	Fußball	10 LU

Tab. 4: Reihenfolge der 16 Ausschnitte.

Zu Beginn eines Durchlaufs wurde den Testpersonen zunächst der Fragebogen ausgeteilt. Bevor es einige Erklärungen zu dem nachfolgendem Versuch gab, wurde als erstes die Abhörlautstärke des Fernsehers anhand eines kurzen Ausschnittes der SWR Nachrichtensendung „Landesschau Baden-Württemberg“¹⁵⁰ eingestellt, so dass sich jeder wohlfühlte. Danach wurde jeder Gruppe erklärt, was eine Tonmischung bedeutet. Da sich viele „Laien“ unter einer Tonmischung nichts vorstellen konnten, wurde anhand eines kurzen Filmausschnitts von „TerraX – Stonehenge“¹⁵¹ direkt in einer ProTools Session gezeigt, dass das Wort und die Musik bzw. Geräusche immer getrennt vorhanden sind. So konnte dargestellt werden, wie man das Mischungsverhältnis als Tonmeister selbst „in der Hand hat“. Anhand des gleich laut bleibenden Wortes wurde demonstriert, wie sich die Lautstärke der Musik und der Geräusche verändern kann. Es wurde das Extrem gezeigt, wobei Musik und Geräusch so laut waren, dass die Sprache komplett verdeckt war, sowie das Gegenteil. So hatten die Probanden

¹⁴⁷ Anm. d. Verf.: Der Mac Mini Server ist ein Kompaktcomputer der Firma Apple.

¹⁴⁸ Anm. d. Verf.: Der VLC (Video Lan Client) ist ein Medienplayer für sämtliche Audio- und Videoformate.

¹⁴⁹ siehe Anhang 6.4.1 *Randomizer*

¹⁵⁰ siehe Anhang 6.3.1.1 *Landesschau Baden-Württemberg*

¹⁵¹ siehe Anhang 6.3.1.2 *TerraX – Stonehenge*

eine ungefähre Vorstellung, was es heißt, wenn die Musik zu laut bzw. zu leise ist. Nach dieser kurzen Einführung wurden alle vier Filmausschnitte in der Sendefassung gezeigt. Die Erfahrungen aus einem vorhergehenden Probanden-Testlauf mit Kollegen hatten gezeigt, dass durch den vorher bekannten Inhalt, die Probanden beim Versuch selbst nicht allzu sehr von der Geschichte abgelenkt waren. Damit es keine Probleme mit der Fragestellung für die 16 Ausschnitte gab, wurde nochmals der Ausschnitt aus „TerraX – Stonehenge“ gezeigt. Die Probanden durften zu diesem Ausschnitt alle drei Fragen als Test bewerten. Nun waren alle Probanden aufgeklärt und der Durchgang mit den 16 Beispielen konnte beginnen. Nach den ersten acht Ausschnitten wurde eine kurze Pause von etwa 10 Minuten eingelegt, damit die Probanden sich erholen konnten.

3.2.8 Ergebnisse und Diskussion

Alle Daten wurden mit der Software „Spss Statistics“¹⁵² erfasst und zusammen mit Marianne Blumers, der Medienforscherin des SWRs, ausgewertet. Zuerst wurden die allgemeinen Fragen des Fragebogens analysiert. Die 21 Testpersonen setzten sich aus 10 Männern und 11 Frauen zusammen, das Verhältnis der Geschlechter war also ausgeglichen¹⁵³. Die Frage, wie man selbst das eigene Hörvermögen beurteilen würde, wurde von 66,7 %¹⁵⁴ mit „gut“ beantwortet. Dies bestätigten auch zum großen Teil die Hörtests.

In Bezug auf die Frage, mit welcher Geräteausstattung die Probanden fernsehschauen, wurde festgestellt, dass 57,1 % einen Flachbildfernseher benutzen. Nur 14 bzw. 19 % besitzen zusätzliche Lautsprecher oder eine home cinema Anlage¹⁵⁵. Die Problematik, dass beim Fernsehen die Sprache manchmal nicht zu verstehen ist, bestätigten 66,7 %¹⁵⁶ der Befragten, wobei davon mehr als die Hälfte (57,1 %) die meisten Schwierigkeiten bei Spielfilmen haben¹⁵⁷. Die Beschwerden bezüglich der zu lauten Hintergrundgeräusche liegen nur bei ca. 14 %¹⁵⁸. Eher wird das Problem noch beim Verstehen der Sprache bei Unterhaltungsshows und Livesendungen gesehen. Die Frage nach den Problemen, die bei dem Verstehen der Sprache auftreten, ergab ein ähnliches Ergebnis wie auch in den analy-

¹⁵² Anm. d. Verf.: Spss Statistics ist eine Statistik und Analyse Software der Firma IBM.

¹⁵³ siehe Anhang 6.3.6.1 *Auswertung Fragebogen gesamt Mittelwerte*, S. 1.

¹⁵⁴ ebd.

¹⁵⁵ ebd. S. 2.

¹⁵⁶ ebd. S. 3.

¹⁵⁷ ebd.

¹⁵⁸ ebd. S. 4.

sierten Zuschauerbeschwerden zu erkennen war: Fast 40 %¹⁵⁹ empfinden die Lautstärkeunterschiede zwischen Sprache und Musik in einer Sendung unangenehm, wobei sich in den Zuschauerbeschwerden dagegen nur knapp 19 % über die undeutliche Sprache der Schauspieler beschwerten. Immerhin gaben fast 50 % der Probanden dieses Problem für das Nicht-Verstehen der Sprache an. Die drei Fragen für den Hörversuch wiesen einige Problematiken auf, die in verschiedene Richtungen interpretiert werden können. Die Beispiele waren bewusst so angelegt, dass es keine plakativen Teile gab, in denen das Wort komplett von der Musik oder den Geräuschen verdeckt wurde. Die Frage, ob jedes Wort verstanden wurde, fiel demnach zu großem Teil positiv aus. Nur bei drei Beispielen gab es jeweils zwei Personen, die „trifft weniger zu“ angekreuzt haben, bei zwei Beispielen nur jeweils eine Person. Ansonsten wurde in allen Filmausschnitten alles verstanden. Die Frage nach der Höranstrengung zeigte einige Schwierigkeiten, die die Probanden trotz gezielter Einführung bei den unterschiedlichen Beispielen hatten. Verifiziert wird das Ergebnis der zweiten Frage durch die Antwort der dritten Frage. Ist die Höranstrengung groß, „müsste“ das Lautstärkeverhältnis zwischen Sprache und Musik/Geräusch als unangenehm empfunden werden. Dies wurde in den meisten Fällen auch bestätigt. Vergleicht man beide Ergebnisse, erhält man folgendes Diagramm:

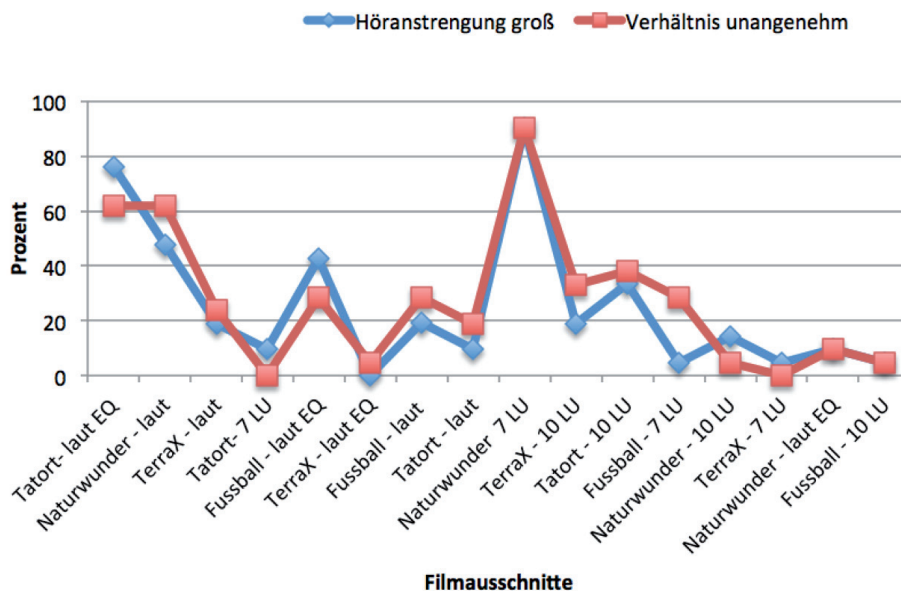


Abb. 25: Vergleich Frage 2 und 3.¹⁶⁰

¹⁵⁹ ebd. S. 4.

¹⁶⁰ siehe Anhang 6.3.6.2 Vergleich Frage 2 mit Frage 3

Insgesamt wurden die Ausschnitte, in denen die Sprache 7 LU Abstand zur IT-Mischung hatte, durchgehend als sehr positiv bewertet. Beim Fußball- und dem TerraX Ausschnitt gaben 95,2 %¹⁶¹ an, dass sie sich bei den 7 LU Beispielen weniger bzw. gar nicht anstrengen mussten, um die Sprache zu verstehen. Beim Tatort – Ausschnitt waren es 90,4 %¹⁶². Ein Ausreißer war allerdings der Ausschnitt aus der „Show der Naturwunder“. Hier gaben 90,4 %¹⁶³ an, dass die Höranstrengung trotz 7 LU Abstand groß war. Dies könnte damit zusammenhängen, dass dieses Beispiel als erstes nach dem zweiten Versuchsblock gezeigt wurde. Eventuell waren die Zuhörer durch die Diskussion, die in der Pause geführt wurde, besonders kritisch und dadurch bei der Beantwortung der ersten Frage nach der Pause noch beeinflusst. Trotzdem könnte man die 7 und 10 LU Versionen als für den Zuschauer angenehmste Mischung sehen.

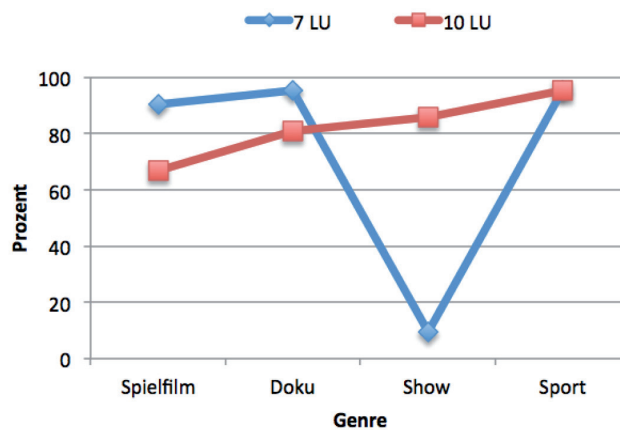


Abb. 26: Bewertung der Mischvarianten 7 und 10 LU „Höranstrengung trifft nicht zu“.¹⁶⁴

Nicht ganz eindeutig war die Aufteilung beim Fußball-Ausschnitt: Obwohl beim Beispiel „laut + EQ“ 42,8 %¹⁶⁵ angaben, dass die Höranstrengung groß war, empfanden trotzdem fast 30 %¹⁶⁶ das Verhältnis von Sprache und Atmo als angenehm. Bekräftigt hat dies die Diskussion, die jeweils nach dem Hörtest stattfand. Viele Probanden hören gerne mehr Atmo bei Fußball-Spielen, um das Gefühl zu haben live im Stadion dabei zu sein.

Des Weiteren gab es eine leichte Tendenz in Bezug auf das Alter. Werden die Probanden in

¹⁶¹ siehe Anhang 6.3.6.3 Zusammenfassung 4 Ausschnitte, S. 1f.

¹⁶² ebd.

¹⁶³ ebd.

¹⁶⁴ siehe Anhang 6.3.6.4 Vergleich 7 LU und 10 LU Variante

¹⁶⁵ siehe Anhang 6.3.6.3 Zusammenfassung 4 Ausschnitte, S. 1f.

¹⁶⁶ ebd.

unter 50 und über 50 Jahre geteilt, werden die ersten acht Beispiele von den über 50jährigen schlechter bewertet als von den unter 50jährigen. Die älteren Personen mussten sich also in der ersten Hälfte der Beispiele weitgehend mehr anstrengen, um die Sprache zu verstehen.

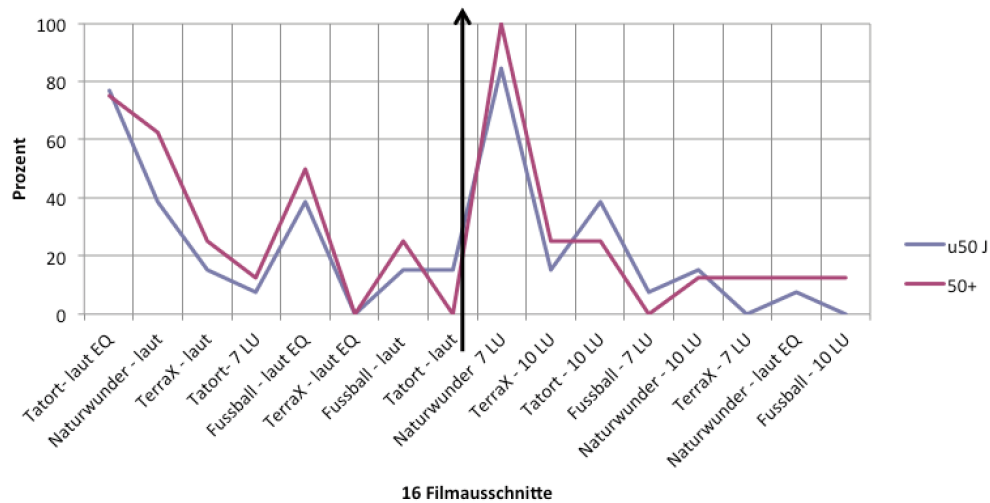


Abb. 27: Aufteilung der Probanden in unter 50 und über 50 Jahre bei der Bewertung der Höranstrengung.¹⁶⁷

Die auffallende Tendenz in Bezug auf die erste Hälfte der Beispiele könnte auch damit zusammenhängen, dass der Randomizer, der für die Zufallsreihe eingesetzt wurde, eine nicht gleichmäßig verteilte Abfolge der Beispiele festlegte. So waren die ersten acht Ausschnitte, bis auf einen, fast nur „laut“ oder „laut + EQ“ Fassungen. Dies könnte eventuell das Ergebnis beeinflusst haben, wobei es bei einer anderen Zufallsfolge fraglich ist, ob die Probanden dann anders entschieden hätten.

In der Diskussion nach dem Hörvergleich wurde außerdem deutlich, dass die Höranstrengung durch den schlechten Originalton und das zusätzlich undeutliche Sprechen des Schauspielers Felix Klare im Tatort-Ausschnitt unabhängig von der Musik groß war. Das Beispiel „Tatort – laut EQ“ wurde demnach zu 76,1 %¹⁶⁸ als schwer zu verstehen eingestuft. Auch die Szene aus der „Show der Naturwunder“ war für die meisten bezüglich des Verhältnisses zwischen Sprache und Musik/Atmo nicht optimal, da schon der Originalton durch die lauten Experiment-Geräusche stark gestört war.

¹⁶⁷ siehe Anhang 6.3.6.5 Auswertung Höranstrengung in Bezug auf unter 50 und über 50 Jahre

¹⁶⁸ siehe Anhang 6.3.6.1 Auswertung Fragebogen gesamt Mittelwerte, S. 7

3.2.9 Lösungsansatz für die Mischung

Die Ergebnisse des Hörvergleichs haben gezeigt, dass das Thema Mischung ein Aspekt ist, der von zwei wichtigen Gesichtspunkten abhängt. Zum einen ist das Lautstärkeverhältnis zwischen Sprache und Musik/Geräusch ausschlaggebend, zum anderen die Qualität des Originaltons. Die Probanden erkannten Unterschiede in den einzelnen Mischungen, von denen zum großen Teil die beiden Varianten mit 7 LU und 10 LU Abstand durchgehend als positiv bewertet wurden. Für eine Production Guideline sollen diese beiden Werte der erste Versuch sein, Grenzwerte für das Verhältnis zwischen Sprache und Musik/Geräusch festzulegen. Ist die Sprache also mindestens 7 LU lauter als die Musik/Geräusch, kann von einer guten Sprachverständlichkeit ausgegangen werden. Die angegebene Toleranz bis zu 10 LU Abstand zwischen Sprache und Musik/Geräusch soll dem Tonmeister einen Spielraum bieten, in dem die Sprachverständlichkeit gegeben sein muss. Der endgültige Abstand hängt dann von der Beschaffenheit des Originaltons und von der Art der Musik/Geräusche ab. Diesen sollte ein Tonmeister aber in der Mischung selbst beurteilen dürfen, da er sonst in seiner Kreativität sehr eingeschränkt wäre. Die unterschiedlichen Ergebnisse in Bezug auf die Fassungen „laut“ und „laut + EQ“ beschreiben zusätzlich, dass der Tonmeister aus technischer Sicht extrem in die Mischung eingreifen kann. Für eine Production Guideline können dadurch Mischungen für altersschwerhörige Zuschauer angepasst werden, ohne dass der Normalhörende die z.B. gefilterte Musik als unangenehm empfindet.

Zusätzlich müssen aber auch Probleme betrachtet werden, die die Sprache unabhängig von der Lautstärke beeinflussen. Die Qualität des Originaltons spielt demnach genauso eine Rolle wie die Auswahl der Instrumente für die Musik. So stört z.B. ein lautes Melodie-Instrument, das seinen Frequenzbereich im Formantbereich der Sprache besitzt, mehr als eine gleichmäßige Klangfläche. Bezieht man sich auf die Empfehlungen, die die BBC in den „best practise tips“ herausgegeben hat, darf man auch den rhythmischen Aspekt der Musik nicht ignorieren.¹⁶⁹ Für die Entwicklung einer Production Guideline für Sprachverständlichkeit sind dies nicht zu vernachlässigende Aspekte, deren Weiterverfolgung den zeitlichen Rahmen dieser Arbeit überschreiten würde.

¹⁶⁹ vgl. BBC, „Hearing Impaired Audiences“, <http://www.bbc.co.uk/guidelines/editorialguidelines/page/guidance-hearing-full> (08.01.2014)

3.3 Abhörsituation während der Produktion

Ein weiterer Aspekt, warum die Tonqualität im Fernsehen nicht immer optimal ist, kann an der Hörsituation in Produktionsregien liegen. Mischstudios der Rundfunkanstalten und Filmstudios sind auf einem sehr hohen akustischen und technischen Standard, so dass Mischungsverhältnisse vom Tonmeister gut beurteilt werden können. Kritisch wird es allerdings dann, wenn Redakteure, Regisseure oder Produzenten nicht am besten Abhörpunkt wie der Tonmeister sitzen und eine Mischung beurteilen müssen. Hier kann es durch die andere Entfernung zum Abhörlautsprecher zu einer unterschiedlichen Bewertung des Sprache/Musikverhältnisses kommen. In Folgendem wird beschrieben, wie sich die Wahrnehmung der Abhörsituation verändern kann und wie man diese in Bezug auf die Sprachverständlichkeit verbessern könnte.

3.3.1 Produktionsregie

Die Tonregien, in denen Fernsehfilme gemischt werden, sind meistens sehr ähnlich aufgebaut. Das Abhörsystem ist üblicherweise als 5.1. System mit qualitativ hochwertigen Lautsprechern angelegt. Damit können auch Stereo – Mischungen durch „Nicht-Beschalten“ der Center- und der Surroundlautsprecher in guter Qualität abgehört werden. Üblicherweise sind zusätzlich kleine Nahfeldmonitore vorhanden, die in der Dynamik und im Frequenzbereich eingengt sind, um die Mischung auch in minderer Qualität beurteilen zu können. Obwohl die meisten professionellen Studios akustisch so gestaltet sind, dass sie ein möglichst neutrales Klangbild abgeben, ermöglicht die Positionierung der Nahfeldmonitore im Direktfeld einen vom Raum weniger beeinflussten Klang.¹⁷⁰

Der Platz für den Regisseur, Redakteur oder Cutter ist meist mit Abstand hinter dem des Tonmeisters aufgebaut. Hier beurteilen die anwesenden Produktionsmitglieder, ob der Klang und die Mischung ihren Wünschen und Vorstellungen entsprechen. Durch die unterschiedliche Hörposition kommt es nicht selten zu unterschiedlichen Beurteilungen zwischen Tonmeister und z.B. der Redaktion. Die häufigsten Diskussionen beziehen sich auf das Mischungsverhältnis zwischen Sprache und Musik/Atmo. Deshalb wurde beim Norddeutschen Rundfunk in Hamburg im „Synchronstudio 1“ ein Mischlabor eingerichtet, mit dem herausgefunden werden sollte, weswegen es zu diesen Diskussionen kommt und wie man diese in Zukunft vermeiden könnte.

¹⁷⁰ vgl. Raffaseder, Hannes: Audiodesign, S. 178f.

Die folgenden Versuche beruhen auf Erkenntnissen, die mit Diplom-Tonmeister Askan Siegfried (NDR) sowie Toningenieur Erich Ebert (SWR) in Hamburg beim Norddeutschen Rundfunk ermittelt wurden. Aufgrund dieser subjektiven Beurteilung entspricht folgende Untersuchung keinen wissenschaftlichen Qualitätskriterien. Dennoch geben die Ergebnisse deutliche Hinweise, die durch die langjährigen Erfahrungswerte der Kollegen bei der Entwicklung einer *Production Guideline* zu einer Verbesserung der Sprachverständlichkeit beitragen können.

3.3.2 Synchronstudio/NDR - Problematik

Am Beispiel des „Synchronstudio 1“ beim Norddeutschen Rundfunk in Hamburg, wurden einige Vorversuche bezüglich der Wahrnehmung an verschiedenen Hörpositionen durchgeführt. Das Synchronstudio ist folgendermaßen eingerichtet:

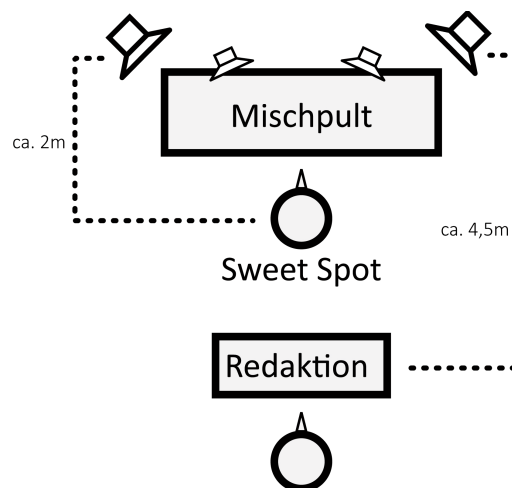


Abb. 28: Synchronstudio 1 NDR (eigene Grafik).

Ausgestattet ist das Studio mit fünf aktiven Geithain (Serie RL 901K) Lautsprechern und zwei Klein und Hummel Nahfeldmonitoren (Serie M52)¹⁷¹, die auf der Kante des Mischpultes positioniert sind. Da die Mischbeispiele ausschließlich in Stereo vorbereitet waren, wurden der Center- und die Surround Lautsprecher in den Versuchen nicht beschaltet und in der Grafik vernachlässigt. Ca. 2 Meter hinter dem Sweet Spot ist der Platz für die Redaktion, die Entfernung dort zu den Regielautsprechern beträgt etwa 4,5 Meter.

Für die Untersuchung wurden mehrere kurze Filmausschnitte und längere Passagen aus

¹⁷¹ siehe Anhang 6.1.4 Email Timm Hübner, 01.10.2013

unterschiedlichen Genres abgespielt. Ziel war es herauszufinden, in wieweit sich die Wahrnehmung der Sprache über die unterschiedlichen Abhör-Systeme ändert und ob es einen Zusammenhang zwischen Abhörsituation während der Produktion und den Zuschauerbeschwerden gibt. Zunächst wurde die Lautstärke der einzelnen Lautsprecher-Systeme subjektiv eingestellt. Dazu wurde ein Ausschnitt der ARD „Tagesschau“ abgespielt und die als angenehm empfundene Lautstärke danach festgelegt. Gemessen wurden die Abhörpegel mithilfe des Brüel und Kjaer Schallpegelmessers „Mediator“ und lieferte folgende Ergebnisse (A –bewertet):

Gemessen am Sweet Spot ergab der Pegel des Geithain Abhörsystems 58 dB, gemessen am Redaktionsplatz 53 dB. Der Abstand der Geithain Lautsprecher vom Sweet Spot zum Redaktionsplatz ergibt also einen Pegelunterschied von 5dB. Diese Erkenntnis hat die Erfahrungen beim Hören der Filmausschnitte unterstützt. Die K&H am Pult ergaben am Sweet Spot 60 dB.

Somit war zu erwarten, dass die Mischungen jeweils an den beiden Abhörpunkten „Sweet Spot“ und „Redaktionsplatz“ unterschiedlich klingen. Nicht nur die Dynamik und das Klangbild änderten sich, sondern auch die Abbildung der Sprache. So war zu erkennen, dass die Sprachverständlichkeit am Redaktionsplatz bei Filmausschnitten mit Musik wesentlich „günstiger“ ausfiel als am Sweet Spot. Dies könnte mit der Lautstärkeabnahme von 5 dB zusammenhängen, da das Signal aus links und rechts in 4,5 Metern Entfernung indirekter, das gleichphasige Mittensignal aber durch die Addition von L+R direkter wirkt. Bis zu den Ohren des Redakteurs verschwimmen so die Stereo Signale (Musik), weswegen vermutlich am Redaktionsplatz die Beurteilung von kritischen Stellen in Bezug auf die Sprachverständlichkeit wesentlich schwerer ist. Nimmt man also an, dass am Redaktionsplatz die Musik tendenziell indirekter wahrgenommen wird als die Sprache, könnte es den Fall geben, dass der Redakteur mehr Musik dazu gemischt haben möchte. An seinem Platz wird bis zu einem gewissen Grad die Sprache noch nicht so gestört wie es am Sweet Spot Platz gehört wird.

3.3.3 Empfehlung für Produktionsregien

Um das Klangbild auch außerhalb des Sweet Spots auch am Redaktionsplatz dahinter präsent und deutlich wahrnehmen zu können, wurden dort zusätzlich zwei Lautsprecher der Marke K&H ¹⁷² aufgestellt. Diese wurden auf einen Abhörpegel von 51 dB eingestellt, der zusammen mit den Regielautsprechern kaum zu hören war. Das Ergebnis war dennoch eine wesentliche Verbesserung der Präsenz der Mischung. Der Pegelabfall von 5 dB durch die Entfernung zur Regieabhöre wurde dadurch korrigiert. Zusätzlich wurden die Nahfeldmonitore mit einem Delay von 12 ms verzögert und so die Laufzeit zu den Hauptlautsprechern angeglichen. Damit bleibt die Vorne-Ortung erhalten und die Direktheit des Klangbildes entsprach dem Sweetspot. Dabei ist allerdings zu beachten, dass die Lautstärke der zusätzlichen Lautsprecher nur eingeschränkt geändert werden kann. Ansonsten besteht die Gefahr, dass die Produktion-Abhöranlage durch einen zu hohen Abhörpegel den Tonmeister bei der Arbeit stört. Außerdem sollte der Abhörpegel der zusätzlichen Lautsprecher an den Abhörpegel der Regielautsprecher gekoppelt sein, um ebenfalls die bestmöglichen Voraussetzungen für das Arbeiten des Tonmeisters zu schaffen.

Des Weiteren wurde durch die Versuche im NDR Synchronstudio festgestellt, dass ein Consumer-Fernseher in Form eines Flachbildfernsehers die beste Referenz darstellt. In den meisten, der heute angebotenen Flachbildschirm-Fernsehapparaten sind kleine, oft minderwertige Lautsprecher eingebaut, die entweder nach unten oder nach hinten abstrahlen. Der Grund ist der nicht vorhandene Platz, da die Fernseher immer flacher werden und seitlich, nach vorne abgestrahlte Lautsprecher den Platz für das Bild einschränken würden. Dies führt dazu, dass die Fernseher an sich bereits keine guten tontechnischen Voraussetzungen bieten. Werden sie dann zu Hause auch noch ungünstig aufgestellt, ist der beim Zuschauer ankommende Ton stark in der Qualität beeinträchtigt. Für Produktionsregien ist somit zu überlegen, wie man als zusätzliche „Referenzabhöre“ einen Consumer Fernseher integriert. So könnten kritische Stellen, die dem Tonmeister auffallen, nochmals mit der Redaktion oder dem Regisseur überprüft werden.

¹⁷² Anm. d. Verf.: Klein & Hummel M52

4 Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurde die Thematik „Sprachverständlichkeit im Fernsehen“ behandelt und in Bezug auf die Erstellung einer Guideline für die Produktion von Sendungen untersucht. Durch eine Analyse von insgesamt 209 Zuschauerbeschwerden, die in Zusammenarbeit mit der ARD und dem ZDF zur Verfügung gestellt wurden, kristallisierten sich die Probleme „undeutlich sprechende Schauspieler“, sowie „zu laut gemischte Musik unter der Sprache“ heraus. Deshalb wurde auf die beiden Themengebiete „Sprechwissenschaften“ und „auditive Wahrnehmung“ im Kapitel 2 dieser Arbeit eingegangen. Innerhalb der Sprechwissenschaften geht es zunächst um den Begriff ästhetische Kommunikation, sowie Sprechbildung und Sprecherziehung von Schauspielern. Das Thema auditive Wahrnehmung befasst sich mit dem Problem der Hörfläche des menschlichen Gehörs, dem Frequenzspektrum der Sprache und auch dem Problem der Verdeckung. Letzteres ist besonders wichtig, da sich viele Zuschauer über die zu laute Musik in Sendungen beschwerten.

Im Hauptteil der Arbeit wurde ein Hörvergleich zum Thema „zu laut gemischte Musik bzw. Atmo“ mit 21 Probanden beim SWR in Stuttgart durchgeführt. Dazu wurden vier Film-Ausschnitte unterschiedlicher Genres ausgewählt, die jeweils mit vier unterschiedlichen Mischungsvarianten unterlegt wurden. Diese Mischungen unterschieden sich im Lautstärkeverhältnis zwischen Sprache und Musik und mussten anhand eines Fragebogens von den Testpersonen bewertet werden.

Außerdem wurde beim NDR in Hamburg ein „Mischlabor“ eingerichtet, in dem einige Untersuchungen zum Thema Abhörsituation in Produktionsregien gemacht wurden.

Die Ergebnisse der Untersuchungen bestätigen, dass das Thema Sprachverständlichkeit im Fernsehen von vielen unterschiedlichen Aspekten abhängig ist, die nicht ausschließlich durch eine veränderte Mischung seitens des Tonmeisters optimiert werden können. Bevor man Empfehlungen für die Mischung aussprechen kann, muss zunächst der Originalton technisch so gut wie möglich aufgenommen werden. Dabei geht es um Fragen wie den Drehort, das richtige Equipment und auch die sprecherische Leistung der Schauspieler. Diese Punkte betreffen nicht nur das Sounddepartment, sondern müssen von Regisseur, Redakteur und Produzent ebenfalls bedacht und überprüft werden. Wenn diese Punkte bestmöglich erfüllt werden, sind für die Mischung gute Voraussetzungen gegeben. Trotzdem gibt es vor allem für die Postproduktion Empfehlungen, die aus den Ergebnissen dieser Arbeit hervorgehen.

Abstand zwischen Sprache und Musik/Geräusche:

Für das Lautstärkeverhältnis zwischen Sprache und Musik/Geräusch wird ein Abstand von mindestens 7 LU empfohlen. Ist die Musik/Geräusch um diesen Wert leiser, kann davon ausgegangen werden, dass die Sprache nicht von der Musik oder den Hintergrundgeräuschen zu sehr verdeckt wird. Abhängig ist dies auch von der Wahl und Art der Musik sowie der Platzierung von Soundeffekten. Ein Musikstück mit einem Melodieinstrument kann die Sprache eher stören, als eine Klangfläche. Ebenso kann es Szenen geben, in denen ein gut platzierter Sound lauter dazu gemischt werden kann. Dies funktioniert aber nur in Sprachpausen, die nicht zu kurz sein dürfen. Werden Geräusche oder Soundelemente über die Sprache gelegt, können Konsonanten oder Vokale verdeckt werden. Besonders bei kritischen Stellen muss also der Abstand zwischen Sprache und Musik/Geräusch überprüft werden.

Referenzabhöre:

Für Produktionsregien empfiehlt es sich, eine Referenzabhöre zur Verfügung zu stellen. Zum einen muss der Abhörplatz des Redakteurs/Regisseurs ausreichend mit Zusatzlautsprechern unterstützt werden. Mit kleinen Nahfeldmonitoren kann die Direktheit und das ausgeglichene Klangbild optimiert werden. Zum anderen wird empfohlen einen Consumer-Fernseher als Abhöre bereitzustellen. Dieser spiegelt am besten die Abhörsituation in einem „Standard Wohnzimmer“ wieder und dient als beste Kontrolle bei problematischen Stellen. Die Schwierigkeit besteht dennoch darin, dass es bisher keine technische und somit objektive Möglichkeit gibt, die Sprachverständlichkeit einer Produktion nachzuweisen. Auch kommt hinzu, dass auf einige Aspekte, wie das Endgerät oder die eventuelle Hörverminderung der Zuschauer, kein Einfluss seitens der Produktion genommen werden kann. Das Problem Sprachverständlichkeit ist unter Tonmeistern, die für den Rundfunk arbeiten, ein brisantes Thema. Trotzdem können die verschiedenen Aspekte der Sprachverständlichkeit nicht nur mithilfe einer offiziellen technischen Richtlinie für Techniker behoben werden. Das Verständnis von „guter Sprachverständlichkeit“ ist ein Begriff, der bisher nur subjektiv beurteilt werden kann. Deshalb müssen Regisseure, Redakteure und Produzenten sensibilisiert werden, damit die Probleme, die durch qualitativ schlechten Originalton, undeutlich artikulierende Schauspieler oder ungünstig gewählte Hintergrundmusik entstehen, weitgehend vermieden werden. Dies stellt eine große Herausforderung dar, da sich dadurch viele

Redakteure und Regisseure in ihrer kreativen Arbeit eingeschränkt fühlen. Meist trifft der Regisseur die Entscheidung, an welchem Ort gedreht wird oder wie der Schauspieler sprachlich agieren soll. Verlangt er, dass der Schauspieler wie im genannten Polizeiruf „Der Tod macht Engel aus allen“, textlich improvisieren soll, sollte trotzdem klar sein, was dies für Auswirkungen auf die Tonqualität haben kann. Oft ist den Regisseuren oder Redakteuren einer Sendung nicht bewusst, welche Auswirkungen manche unkonventionelle Ideen auf den Ton haben können. So ist der Zuschauer auch oft von neuen Sounddesign – Effekten, die jede Bewegung im Bild oder den Bildschnitt selbst unterstützen überfordert, obwohl das in vielen Sendungen als Modernisierung gilt. Auch bei der Untermalung von Beiträgen mit Musik fragt sich der Großteil der Zuschauer, wieso Musik unter Erklärungen notwendig ist. Nicht vergessen werden darf das Durchschnittsalter der Zuschauer bei den öffentlich-rechtlichen Sendern in Deutschland. Zwar möchten diese jüngere Zuschauer gewinnen, zunächst stellt sich aber die Frage „wie kann ein Sender jünger werden“? Ist es ausschließlich der Ton, der mit seinen lautmalerischen Effekten jüngeres Publikum anzieht oder sollte eventuell erst einmal der Inhalt der Sendungen hinterfragt werden bevor in Sounddesign investiert wird, wenn es um altersrelevante Themen geht. Natürlich soll nicht ausschließlich für ein hörvermindertes Publikum gemischt werden, aber dennoch kann weit in die Mischung eingegriffen werden, damit das Hörerlebnis für z.B. Altersschwerhörige verbessert wird, ohne dass sich das jüngere Publikum beschwert. Zu überdenken wäre auch der Vorschlag, die Mischungen auf das Genre und den Inhalt zu optimieren. So könnte man bei Sendungen, die inhaltlich eher das ältere Publikum ansprechen, dementsprechend die Mischung für eventuell hörverminderte Zuschauer ausrichten und bei Beiträgen für Jüngere das Sounddesign entsprechend anpassen. Hinter dem Aspekt, dass unsere Gesellschaft immer älter wird, ist der Gedanke nicht abwegig, dass das Thema Sprachverständlichkeit in naher Zukunft auch in die „*Technischen Richtlinien - HDTV zur Herstellung von Fernsehproduktionen für ARD, ZDF und ORF*“ aufgenommen wird und dadurch als technische Vorgabe bindend wird.

5 Literaturverzeichnis

Aderhold, Egon

Sprecherziehung des Schauspielers. 4. überarbeitete Auflage Berlin: Henschel Verlag, 1962

Barthel, Henner (Hg.)

Zum Wissenschaftsverständnis der Sprechwissenschaft. München: Ernst Reinhardt, GmbH & Co KG, 2003

British Broadcast Corporation

Hearing Impaired Audiences.

<http://www.bbc.co.uk/guidelines/editorialguidelines/page/guidance-hearing-full>
(08.01.2014)

Chion, Michel

Audiovision - Sound on Screen. New York: Columbia University Press, 1994

Cohen, Danny

Clear Sound – best practice tips.

<http://www.bbc.co.uk/academy/production/article/art20130702112135255>
(13.01.2014)

Dickreiter, Michael

Handbuch der Tonstudioteknik. Bd. 1, 6. verbesserte Auflage München: K.G. Saur Verlag KG, 1997

Handbuch der Tonstudioteknik. Bd. 2, 6. verbesserte Auflage München: K.G. Saur Verlag KG, 1997

Dittmann, Klaus

Umfrage. <http://www.schwerhoerigen-netz.de/MAIN/termine.asp?inhalt=2013/UMFRAGE/umfrage> (25.02.2014)

Fiukowski, Heinz

Sprecherzieherisches Elementarbuch. 7. neu bearbeitete Auflage. Tübingen: Max Niemeyer Verlag GmbH, 2004

Fleischer, Gerald

Gut Hören - Heute und Morgen. Heidelberg: Median - Verlag von Killisch-Horn GmbH, 2000

Goldstein, E. Bruce, Manfred Ritter und Gabriele Herbst

Wahrnehmungspsychologie. 2. dt. Aufl. Spektrum Lehrbuch Heidelberg [u.a.]: Spektrum, Akad. Verl., 2002

Görne, Thomas

Tontechnik. 3., neu bearb. Aufl. München: Hanser, 2011

Gutenberg, Norbert

Grundlagenstudien zu Sprechwissenschaft und Sprecherziehung: Kategorien - Systematik - Programm. Bd. 615. Göppinger Arbeiten zur Germanistik Göppingen: Kümmerle Verlag, 1994

Einführung in Sprechwissenschaft und Sprecherziehung. Frankfurt am Main: Peter Lang GmbH, 2001

Habermann, Günther

Stimme und Sprache. 3. unveränd. Auflage Stuttgart: Georg Thieme Verlag, 2001

Hakkarainen, Heikki J.

Phonetik des Deutschen. München: Wilhelm Fink Verlag GmbH & Co. KG, 1995

Herzog, Uwe

Das Sprecherhandbuch: Ausbildung und Praxis bei Film, TV, Funk Und Werbung. Kindle, 2007

Institut für Rundfunktechnik

Technische Richtlinien – HDTV.

<http://www.irt.de/de/publikationen/technische-richtlinien.html>

(08.01.2014)

kha/dpa

Tonprobleme beim „Polizeiruf“.

<http://www.spiegel.de/kultur/tv/ton-bei-polizeiruf-war-aus-dramaturgischengruenden-schlecht-a-911204.html>

(19.09.2013)

Lazarus Hans, Charlotte A. Sust, Rita Steckel, Marko Kulka und Patrick Kurtz

Akustische Grundlagen sprachlicher Kommunikation. Berlin, Heidelberg und New York: Springer, 2007

Lemke, Siegrun (Hrsg.)

Sprechwissenschaft/ Sprecherziehung: Ein Lehr- und Übungsbuch. Band 4. 2., überarbeitete und ergänzte Auflage Frankfurt am Main: Peter Lang GmbH, 2012

Marschall-Pompino, Bernd

Einführung in die Phonetik. 3. Aufl. Berlin: Walter de Gruyter GmbH & Co. KG, 2009

Nikula, Henrik

Der literarische Text - Eine Fiktion. Tübingen: Narr Francke Attempto Verlag GmbH & Co. KG, 2012

Polenz, Peter von

Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart. Bd. 3. Berlin: Walter de Gruyter GmbH & Co, 1999

Raffaseder, Hannes

Audiodesign. München: Carl Hanser Verlag, 2010

Seidler, Herbert

Grundfragen einer Wissenschaft von der Sprachkunst. München: Wilhelm Fink Verlag, 1978

Spitzer, Manfred

Musik im Kopf: Hören, Musizieren, Verstehen und Erleben im Neuronalen Netzwerk.
6. Nachdr. der 1. Aufl. Stuttgart: Schattauer, 2006

Sundberg, Johan

Die Wissenschaft von der Singstimme. Bd. 86. Orpheus-Schriftenreihe zu
Grundfragen der Musik. Bonn: Orpheus-Verlag GmbH, 1997

Weinschenk-Pabst, Marita

Grundlagen der Sprechwissenschaft und Sprecherziehung. 2. Auflage München: Ernst
Reinhardt, GmbH & Co KG, 2011

Zwicker, Eberhard

Das Ohr als Nachrichtenempfänger. Stuttgart: S. Hirzel Verlag, 1967
Psychoakustik Berlin Heidelberg: Springer -Verlag, 1982

6 Anhang

(siehe DVD)

6.1 Korrespondenz

6.1.1 Email Barbara Bernhard vom 27.09.2013

6.1.2 Email Askan Siegfried vom 05.09.2013

6.1.3 Email Peter Mechtler vom 05.11.2013

6.1.4 Email Timm Hübner vom 01.10.2013

6.1.5 Email Dr. Walker vom 26.11.2013

6.1.6 Email Erich Ebert vom 05.02.2014

6.2 Zuschauerbeschwerden

6.2.1 Zuschauerbeschwerden

6.2.2 ARD Wochenprotokolle

6.2.3 Auswertung Beschwerden

6.3 Hörvergleich

6.3.1 Referenzvideos

6.3.2 Izotope Grafiken

6.3.3 Fragebogen

6.3.4 Werte Mischungsvarianten

6.3.5 Filmausschnitte

6.3.6 Auswertung

6.3.7 Fragebogen der 21 Probanden

6.4 Sonstiges

6.4.1 Randomizer

6.4.2 Press Release – VLV Audibility Project

6.4.3 Durchschnittsalter SWR Zuschauer

6.4.4 Messungen Studio 13