

Besonderheiten der Aufnahmetechnik für Sprachaufnahmen bei Hörbuchproduktionen im Tonstudio

Mikrofonierung und analoge Echtzeitbeeinflussung von Sprachsignalen

Bachelorarbeit im Studiengang Audiovisuelle Medien an der Hochschule der Medien
Stuttgart zur Erlangung des akademischen Grades „Bachelor of Engineering“.

Vorgelegt von: Lucy Duffner

Matrikelnummer: 36787

E-Mail: info@lucyduffner.de

Betreuer: Prof. Oliver Curdt (Erstprüfer)

Prof. Dr. Michael Felten (Zweitprüfer)

Vorgelegt am: 17.02.2022

Ehrenwörtliche Erklärung

Hiermit versichere ich, Lucy Duffner, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „Besonderheiten der Aufnahmetechnik für Sprachaufnahmen bei Hörbuchproduktionen im Tonstudio – Mikrofonierung und Echtzeitbeeinflussung von Sprachsignalen“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.“

Stuttgart, 17.02.2022

Ort, Datum



Unterschrift

Kurzfassung

Die vorliegende Arbeit beschäftigt sich mit Besonderheiten der Aufnahmetechnik, die bei der Realisierung von hochwertigen Sprachaufnahmen für Hörbuchproduktionen im Tonstudio von Bedeutung sind. Wissenschaftlich fundierte Erklärungen qualitätsbestimmender Phänomene belegen die Auswahl und die praktische Anwendung der Aufnahmetechnik bei Hörbuchproduktionen. Dabei wird die erzeugte Sprache als Schallereignis, die Mikrofonierung und die analoge Echtzeitbeeinflussung bei Sprachaufnahmen im Tonstudio untersucht.

Abstract

The subject of the present paper comprehends special features of technical recording equipment which determine the quality of audio book recordings in a sound studio. Academically proven explanations concerning technical phenomena state which recording equipment should be chosen and how voice recordings should be realized in practice. In the course of this, voice generation as a sound event, miking and analogue influencing of recordings in real time, is studied.

Inhaltsverzeichnis

Ehrenwörtliche Erklärung.....	II
Kurzfassung.....	III
Abstract	III
Inhaltsverzeichnis.....	IV
Abbildungsverzeichnis.....	VI
Formelverzeichnis	VIII
Abkürzungsverzeichnis	IX
1. Einführung	1
1.1 Zielsetzung.....	2
1.2 Aufbau der Arbeit	2
1.3 Anmerkungen	3
2. Schall.....	4
2.1 Pegel.....	9
3. Sprachsignal und Spracherzeugung	10
3.1 Akustisches Sprachsignal	10
3.2 Mechanismus der Spracherzeugung.....	11
3.2.1 Quelle-Filter-Modell	14
3.2.2 Formant-Syntesizer	15
3.3 Frequenzanalyse des Sprachsignals.....	16
3.4 Unterschiede zwischen Sprechern	17
4. Mikrofonierung	19
4.1 Richtcharakteristik	20
4.1.1 Bündelungsgrad und Bündelungsmaß	26
4.1.2 Nahbesprechungseffekt.....	27
4.1.3 Umschaltbare Richtcharakteristiken	29
4.1.3.1 Austrian Audio OC818.....	30
4.2 Elektromechanische Wandlung	31
4.3 Frequenzgang und Übertragungsbereich	35
4.4 Übertragungsfaktor und Übertragungsmaß	39
4.5 Nennimpedanz und Nennabschlussimpedanz	40
4.6 Störpegel	41
4.7 Impulsverhalten	44

4.8	Maßnahmen gegen Störschall	45
4.9	Mikrofon positionieren.....	47
4.9.1	Kammfiltereffekt	49
5.	Analoge Echtzeitbeeinflussung von Sprachsignalen	51
5.1	Vorverstärkung.....	55
5.1.1	Mikrofonvorverstärker	55
5.1.2	Nichtlineare Verzerrungen und Klirrfaktor	57
5.1.3	Störabstand und Dynamikbereich	60
5.2	Signalbearbeitung	62
5.2.1	Kompression	62
6.	Fazit	67
6.1	Ausblick	69
	Literaturverzeichnis.....	70
	Anhang	X
	1. Interview mit Markus Götze	X
	2. Hörproben von Hörbuchproduktionen.....	XVIII

Abbildungsverzeichnis

Abbildung 1 Teilchenbewegung bei der Ausbreitung einer Schallwelle in Luft	4
Abbildung 2 Schallausbreitung: Kugelwelle mit Abstand r zur Schallquelle und ebene Welle mit Abstand x von der Schallquelle (weiter entfernt)	6
Abbildung 3 Verlauf des Schalldrucks einer Schallwelle im zeitlichen (a) und räumlichen Verlauf (b). Phasenverschiebung zwischen zwei Signalen (c)	7
Abbildung 4 Schwingungsverläufe einer Sinuswelle (a) und eines Rechtecksignals (b) mit den dazugehörigen Amplitudenspektren bis zur 12. Harmonischen	8
Abbildung 5 Wellenform des Wortes Split. Die markierte Stelle wird als /p/ wahrgenommen	11
Abbildung 6 Mechanismus zur Spracherzeugung beim Menschen	12
Abbildung 7 Artikulation von Lauten als schematische Darstellung	13
Abbildung 8 Quelle-Filter-Modell für Spracherzeugung	14
Abbildung 9 Frequenzspektren (Teiltenspektren) von Vokalen und ihre Resonanzfrequenzen	15
Abbildung 10 Blockschaltbild von Sprachsynthesizer mit drei Formanten in Reihenschaltung	16
Abbildung 11 Screenshot aus iZotope RX Connect: Frequenz-Zeit-Diagramm des Satzes „So einfach ist es.“	16
Abbildung 12 Durchschnittliche Frequenzbereiche für unterschiedliche Vokale	17
Abbildung 13 Durchschnittliche Spektren deutscher Sprecher für unterschiedliche Lautstärken. Männliche Sprecher (links) und weibliche Sprecher (rechts)	18
Abbildung 14 Aufbau eines Mikrofons	19
Abbildung 15 Kapselaufbau eines Druckgradientenempfängers mit akustischen Laufzeitgliedern l und Nierencharakteristik. Verschiedene Schalleinfallsrichtungen (a, b, c)	21
Abbildung 16 Schalldruckdifferenzen bei verschiedenen Frequenzen und Frequenzgang des Druckgradienten Dp zwischen der Membranvorderseite (A) und -rückseite (B)	22
Abbildung 17 Pegelverhältnisse aus verschiedenen Schalleinfallsrichtungen von Richtmikrofonen	24

Abbildung 18 Polardiagramme von idealisierten Richtcharakteristiken	25
Abbildung 19 Mikrofonrichtcharakteristiken und ihre Vergrößerungsfaktoren im Vergleich	26
Abbildung 20 Entfernungsabhängiger Frequenzgang einer Superniere beim Nahbesprechungseffekt. x-Achse: Frequenz in Hz, y-Achse: Schalldruckpegel in dB	28
Abbildung 21 Richtcharakteristiken durch die Addition zweier Nieren	29
Abbildung 22 Plug-in zur Einstellung der Richtcharakteristik des OC818 in der DAW	31
Abbildung 23 Schaltkreis einer elektrostatischen Wandlung	33
Abbildung 24 Höhenanhebung des U87 von Neumann mit und ohne Low-Cut	38
Abbildung 25 Kammfiltereffekt im Raum (links) und resultierender Frequenzgang (rechts)	49
Abbildung 26 Kammfiltereffekt vermeiden: geeignete Mikrofonposition für eine Hörbuchaufnahme mit Maßnahmen gegen Störschall.....	50
Abbildung 27 Pegeldiagramm einer hochwertigen Übertragungskette (oben) und einer Übertragungskette mit geringem internem Pegel (unten)	54
Abbildung 28 Eine Sinusspannung wird durch eine verzerrte Kennlinie verformt.....	57
Abbildung 29 Störabstände im Übertragungssystem	61

Formelverzeichnis

Formel 1 Schallintensität I und Schallleistung P_{ak}	5
Formel 2 Wellenlänge λ	6
Formel 3 Schalldruckpegel L_p	9
Formel 4 Kapazität C_0	33
Formel 5 Kapazität C	33
Formel 6 Resonanzfrequenz f_m	36
Formel 7 Grenzfrequenz f_{RC}	37
Formel 8 Betriebsspannung $u_{Betrieb}$	39
Formel 9 Rauschspannung L_{Rausch}	42
Formel 10 Phasenwinkel φ zwischen direktem und reflektiertem Signal	49
Formel 11 Spannungsverhältnis einer Leerlaufspannung U_0 zur Eingangsspannung U_E	52
Formel 12 Klirrfaktor k	58
Formel 13 Klirrdämpfung ak	59
Formel 14 Kompressionsverhältnis R	64

Abkürzungsverzeichnis

A/D-Wandler	Analog-Digital-Wandler
AKG	Akustische und Kino-Geräte GmbH, Gerätehersteller von Audiotechnik
AM	Amplitudenmodulation
CCIR	Comité Consultatif International des Radiocommunications, engl. für internationaler beratender Ausschuss für den Funkdienst
CKR12	Doppelmembrankapsel aus Keramik
CMV3	Kondensatormikrofon des Herstellers Neumann
DAW	Digital Audio Workstation, engl. für Verbund digitaler Geräte zur Audioverarbeitung
DP	Digital Publishers, Verlag mit Sitz in Stuttgart
FET	Feldeffekttransistor-Schaltung
FM	Frequenzmodulation
HF	Hochfrequenzschaltung
NF	Niederfrequenzschaltung
OC818	Mikrofon mit veränderbarer Richtcharakteristik der Herstellers Austrian Audio
RC-Glied	Schaltungen aus ohmschen Widerstand R und Kondensator C
RMS	root mean square, engl. für quadratisches Mittel (u.a. in der Elektrotechnik)
RME Audio	Hersteller von Audiogeräten
RX	Software zur Audiorestauration des Herstellers iZotope
SNR	signal-to-noise-ratio
SPL	sound pressure level, engl. für Schalldruckpegel
THD	Total Harmonic Distortion, engl. für gesamte harmonische Verzerrung
VCA	Voltage Controlled Amplifier

1. Einführung

Seit einigen Jahren zeichnet sich ein starkes Wachstum der Nutzerzahlen auf dem Hörbuchmarkt ab.¹ 2021 haben laut einer Studie der Amazontochter Audible 16 Millionen Deutsche mindestens ein Hörbuch oder -spiel gehört. Das sind 1,6 Millionen mehr Hörer als im Vorjahr.² Laut dem Hörbuchsprecher Stefan Kaminski liegt das an dem imaginären „Kino im Kopf“, das eine vorgelesene Geschichte bei den Hörern erzeugt.³ Damit sich das Hören möglichst immersiv anfühlt, muss Sprache im Tonstudio qualitativ hochwertig aufgenommen werden. Aber wie klingt eine hochwertige Sprachaufnahme für Hörbücher aus dem Tonstudio? Laut Markus Götze, Hörbuchproduzent und langjähriger Inhaber des Tonstudios Tongoetze, hört sich eine gute Hörbuchaufnahme wie folgt an:

„Für mich ist das eine Aufnahme, die trocken ist. Bei Hörbuchaufnahmen höre ich keinen Raum. Sie ist verdichtet und trotzdem noch dynamisch genug. Das heißt sie ist nicht zu Tode komprimiert wie Werbung. Warm und brilliant [...] finde ich klingt ausgewogen. Zwischen den Sätzen und den Sprechpausen höre ich keinen [...] Raum. Da ist für mich eine angenehme Ruhe. Ich höre keine Klicks, keine Schmatzer. Im Ausnahmefall höre ich reduzierte Mundgeräusche, so dass es natürlich klingt.“⁴

Für die praktische Realisierung einer solchen Aufnahme sind die Auswahl geeigneter Audiogeräte, sowie der angemessene Umgang mit Besonderheiten der Aufnahmetechnik, wichtige Voraussetzungen. Wertige Sprachaufnahmen für Hörbücher in der Praxis umzusetzen, erfordert ein Verständnis für theoretische Aspekte der Mikrofonierung und Echtzeitbeeinflussung von Sprache. Aus diesem Grund thematisiert die vorliegende Arbeit Besonderheiten der Aufnahmetechnik bei Sprachaufnahmen für Hörbuchproduktionen im Tonstudio.

¹ Dennis, L.: Hören statt Lesen. Hörbücher, Hörspiele & Co. werden in Deutschland immer beliebter, in: Forschung und Wissen, 22.05.2021 [online] <https://www.forschung-und-wissen.de/nachrichten/oekonomie/hoerbuecher-hoerspiele-und-co-werden-in-deutschland-immer-beliebter-13374997> [05.11.2021]

² Gräfe, Daniel: Start von Bookbeat. Markt für Hörbücher wächst weiter, in: Stuttgarter Zeitung, 23.07.2017 [online] <https://www.stuttgarter-zeitung.de/inhalt.start-von-bookbeat-markt-fuer-hoerbuecher-waechst-weiter.4be4c487-3054-4d3e-b105-cfd27f5a3031.html> [05.11.2021]

³ Weyerer, Marlene: Jetzt gibt's was auf die Ohren, denn: „Hören ist das neue Sehen“, in: Augsburger Allgemeine, 12.06.2021 [online] <https://www.augsburger-allgemeine.de/kultur/Journal/Die-neue-Lust-am-Audio-Jetzt-gibts-was-auf-die-Ohren-denn-Hoeren-ist-das-neue-Sehen-id59862601.html> [05.11.21]

⁴ Götze, Markus: Interview mit Markus Götze, 05.02.2022, siehe Anhang 1.

1.1 Zielsetzung

In dieser Bachelorarbeit soll dargelegt werden, welche Aufnahmetechnik für Sprachaufnahmen bei Hörbuchproduktionen im Tonstudio geeignet ist. Durch einen fundierten Einblick in die Spracherzeugung, die Mikrofonierung eines Sprachsignals und dessen analoge Echtzeitbeeinflussung wird die praktische Realisierung von Sprachaufnahmen für Hörbücher im Tonstudio erklärt. Dabei beeinflussen spezifische Phänomene der Audiotechnik die Qualität von Sprachaufnahmen. Die Theorie hinter diesen Besonderheiten soll die Verwendung spezieller Aufnahmetechnik für den Leser begreiflich machen. Hierzu werden praktische Vorgehensweisen bei Sprachaufnahmen für Hörbücher im Tonstudio aus wissenschaftlich substantiierter Perspektive betrachtet.

1.2 Aufbau der Arbeit

Der erste Teil der Arbeit beschäftigt sich mit Schall, Grundlagen der Spracherzeugung und Eigenschaften von Sprachsignalen. Dieser Abschnitt schließt eine Definition von Schall ein. Anschließend wird die Mikrofonierung für Sprachaufnahmen bei Hörbuchproduktionen untersucht. Der folgende Teil umfasst eine für Hörbuchaufnahmen geeignete analoge Echtzeitbearbeitung des Sprachsignals. Im Mittelpunkt der Arbeit stehen Besonderheiten der Aufnahmetechnik, die es bei der Aufnahme von Sprache zu beachten gilt. Die Theorie der aus den Besonderheiten resultierenden Phänomene wird erklärt. Daraus soll hervorgehen, welche Aufnahmetechnik sich eignet oder wie der Umgang mit der Technik erfolgt. Im zusammenfassenden Fazit wird diese Arbeit durch einen kurzen Ausblick auf Alternativen zur Sprachaufnahme im Tonstudio und die folgenden Arbeitsschritte beendet.

1.3 Anmerkungen

Komplexe Thematiken sollen dem Leser durch Erklärungen aus der Tontechnik, ausgewählte Abbildungen, mathematische Zusammenhänge und Praxisbeispielen verdeutlicht werden. Um die Erklärungen in vollem Umfang nachvollziehen zu können, sind mathematische, technische und naturwissenschaftliche Grundlagen erforderlich. Auf einige Erläuterungen nebensächlicher Thematiken aus der Audiotechnik wird bewusst verzichtet, wenn diese nicht maßgeblich zur Sache der Arbeit beitragen oder sich nicht in den Rahmen der Arbeit einordnen. Eine Ausnahme bildet die Raumakustik. Für die gesamte Arbeit gehe man davon aus, dass alle Aufnahmen in einem für Sprachaufnahmen optimierten Raum im Tonstudio stattfinden.

Die Arbeit nimmt Bezug auf mein praktisches Studiensemester und meiner Tätigkeit im Tonstudio Tongoetze. Die praktischen Kenntnisse werden durch eigene Praxiserfahrungen und ein transkribiertes Interview mit dem langjährigen Tonstudioinhaber Markus Götze, welches im Anhang⁵ zu finden ist, belegt. Im digitalen Anhang⁶, auf dem beigefügten USB-Stick, befinden sich einige Auszüge aus Hörbuchproduktionen, bei denen ich Sprachaufnahmen im Tonstudio Tongoetze praktisch umgesetzt habe. Alle Klangbeurteilungen beziehen sich auf den Durchschnittshörer. Sie sind dennoch subjektiv und können daher vom Geschmack des Lesers abweichen. Die genannten Empfehlungen basieren auf Erfahrungen aus der Hörbuchproduktion.

⁵ Götze 04.02.2022, siehe Anhang 1.

⁶ Tonstudio Tongoetze: Hörproben von Hörbuchproduktionen, 2020-2022, siehe Anhang 2.

2. Schall

Wie bei allen Audioaufnahmen geht es bei Sprachaufnahmen darum Schall aufzunehmen, beziehungsweise zu wandeln. Ein grundlegendes Wissen darüber, wobei es sich bei einem Schallereignis handelt und welche Größen Schall in der Tontechnik charakterisieren, sind tontechnische Grundlagen. Zwar berechnet man in der Aufnahmepraxis keine genauen Werte oder greift auf Formeln zurück, allerdings sollte man sich als Tontechniker mit folgendem auskennen. So wird ein kompetenter Austausch im Audiobereich möglich und auf die Grundlagen aufbauende Funktionsweisen können leichter begriffen werden. Diese Arbeit beschäftigt sich hauptsächlich mit der Sprache als Schallereignis. In diesem Zusammenhang sei erwähnt, dass bei Sprachaufnahmen im Tonstudio die Raumakustik einen maßgeblichen Einfluss auf die Qualität der Aufnahme hat.⁷ Ein umfassender Einblick in raumakustische Phänomene steht allerdings nicht im Rahmen der Arbeit, weil deren Länge ansonsten überhandnähme.

Der Begriff Schall definiert mechanische Schwingungen und Wellen in einem elastischen Medium. Wenn Sprache vom Menschen abgestrahlt wird, breiten sich Schallwellen im Medium Luft aus. Eine Schallwelle ist eine Folge von Luftverdichtungen und -verdünnungen, die sich im Raum fortpflanzt. Die Ausbreitungsgeschwindigkeit der Schallwelle ist die Schallgeschwindigkeit c . Sie ist zwar in geringem Maß von Umgebungsbedingungen beeinflusst, wird aber üblicherweise als konstant wahrgenommen. Bei einer Raumtemperatur von 20 °C liegt die Schallgeschwindigkeit bei einem Luftdruck von 1013 hPa unter trockenen Luftbedingungen bei 344 m/s.⁸

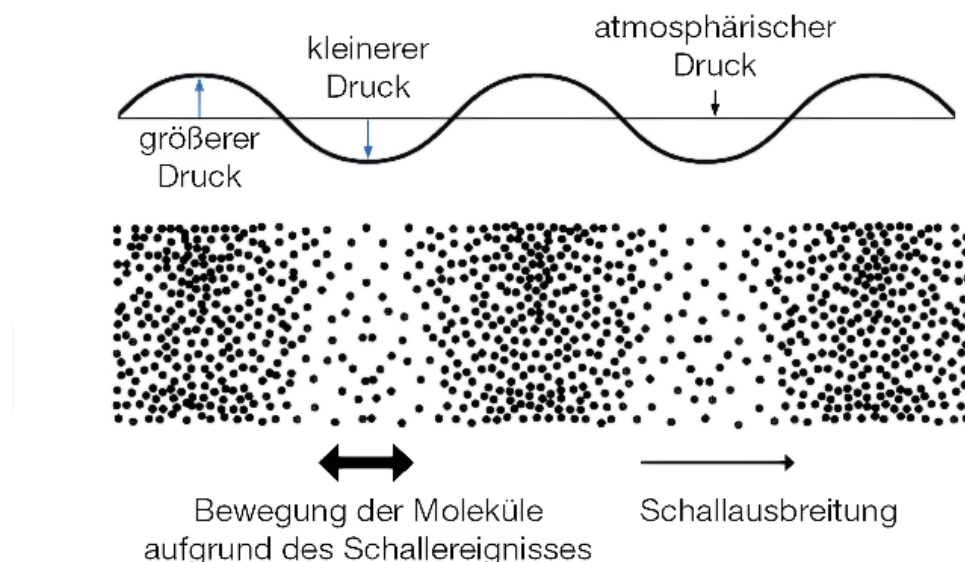


Abbildung 1 Teilchenbewegung bei der Ausbreitung einer Schallwelle in Luft ⁹

⁷ Götze: 05.02.2022, siehe Anhang 1.

⁸ Görne, Thomas: Mikrofone in Theorie und Praxis, 6. Auflage, Aachen: Elektor 2002.

⁹ inVENTer GmbH: Geräuschentwicklung und Schall bei der Lüftung. Grundlagen und Definition von Schall, in: inVENTer, 2022 [online], https://www.inventer.de/wp-content/uploads/2021/03/csm_Schalldruck-und-Uebertragung-der-Molekuele_01d031e2f0.jpg [04.02.2022]

Die kurzfristigen Luftdruckschwankungen entstehen durch schnelles, periodisches Hin- und Herbewegen von Luftteilchen in sehr geringem Ausmaß. Die Luftteilchen bewegen sich um ihre Ruhelage, sodass ihre gemittelte Position unverändert bleibt.¹⁰ Die Schallschnelle v ist die Geschwindigkeit der Teilchen bei dieser Bewegung. Je stärker das Schallsignal ist, desto größer ist die Schallschnelle. Bei sehr lauten Signalen kann sie bis zu einige Millimetern pro Sekunde erreichen.¹¹ In der Regel wird die Schallstärke durch den Schalldruck p angegeben. Dieser gleicht der mittleren Luftdruckschwankung und wird in Pascal (Pa) angegeben. Setzt man Schalldruck und -schnelle in ein Verhältnis, ergibt sich der Wellenwiderstand Z . Er wird auch als Schallimpedanz bezeichnet und ist von der Form der Schallwelle abhängig. Mit zunehmender Entfernung von der Schallquelle nimmt der Schalldruck ab, weil sich die Schallenergie bei der Wellenausbreitung über eine immer größere Fläche verteilt. Nach dem Abstandsgesetz führt eine Verdopplung eines Abstandes zu einer Schallquelle zu einer Halbierung des Schalldrucks.¹²

Schall breitet sich in der Nähe der Schallquelle kugelförmig aus. Die Krümmung der Schallwellen lässt mit zunehmender Entfernung zur Quelle nach. Man unterscheidet die Kugelwelle im Nahfeld von der ebenen Welle im Fernfeld. Die Druckdifferenz zwischen zwei Punkten im Schallfeld wird vom Schalldruckgradienten angegeben.¹³ Zur Beschreibung der Schallquelle eignet sich die Schallintensität I und -leistung P_{ak} . Die Schallintensität ist die Schalleistung, die durch eine Fläche, die senkrecht zur Ausbreitungsrichtung der Schallwelle steht, hindurchströmt. In einer ebenen Welle ist sie das Produkt aus effektivem Schalldruck und -schnelle. Im freien Schallfeld ergibt sich für die Schallintensität in Abhängigkeit von dem Abstand zur kugelförmig schallabstrahlenden Quelle r in m und der Schallleistung P die Schallintensität. Die Schallleistung ist die Summe aller Schallintensitäten und wird in Watt angegeben. Sie bezeichnet die pro Zeitspanne von einer Schallquelle abgegebene Schallenergie und wird über eine Hüllfläche A , die um die Schallquelle gleichmäßig verteilt und zur Schallintensität senkrecht liegend ist, ermittelt. Die Größen stehen in folgendem mathematischen Zusammenhang.¹⁴

Formel 1 Schallintensität I und Schallleistung P_{ak}

$$I = p_{eff} \cdot v_{eff}$$

$$I = \frac{p}{4\pi \cdot r^2}$$

$$P_{ak} = I \cdot A = p \cdot v \cdot A$$

¹⁰ Raabe, Armin/ Holstein, Peter: Akustik und Raumklima. Raumkomfortbewertung und Energieeffizienz, Wiesbaden: Springer Fachmedien 2021.

¹¹ Görne 2002: S.13-14.

¹² Smyrek, Volker: Tontechnik. Für Veranstaltungstechniker in Ausbildung und Praxis, 2. Auflage, Stuttgart: Hirzel 2012.

¹³ Dickreiter, Michael: Mikrofon-Aufnahmetechnik, 3. Auflage, Stuttgart: Hirzel 2003.

¹⁴ Weinzierl, Stefan: Handbuch der Audiotechnik, Berlin/ Heidelberg: Springer 2008.

Die Entfernung einer Schallwelle zu einer Schallquelle steht in Relation zur Wellenlänge λ (Lambda). Erst nach mindestens einer Wellenlänge Abstand zur Schallquelle gelangt man in die ebene Welle im Fernfeld. Die Wellenlänge ist der räumliche Abstand zwischen zwei Druckmaxima einer Welle. Sie wird in Metern gemessen und ist abhängig von Schallgeschwindigkeit und Frequenz.¹⁵ Im hörbaren Frequenzbereich reichen Wellenlängen von 1,5 cm für 20 kHz bis zu 20 m für 16 Hz. Für die Raumakustik und mit ihr in Zusammenhang stehende Phänome ist die frequenzabhängige Wellenlänge entscheidend.

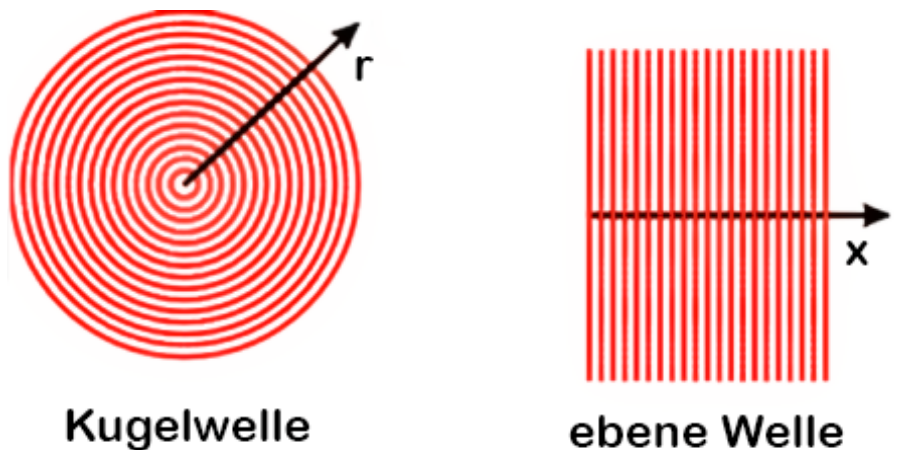


Abbildung 2 Schallausbreitung: Kugelwelle mit Abstand r zur Schallquelle und ebene Welle mit Abstand x von der Schallquelle (weiter entfernt) ¹⁶

Die Frequenz f gibt an, wie oft sich Druckmaxima und -minima abwechseln und wird in Schwingungen pro Sekunde (Hz = Hertz) angegeben. Die Frequenz dient im Audibereich auch als Maß für die Tonhöhe. Der für Menschen hörbare Frequenzbereich erstreckt sich von ungefähr 20 Hz bis 20 kHz. Der Kehrwert der Frequenz ist die Periodendauer T . Sie gibt die Zeit an, die zwischen zwei gleichen Druckzuständen verstreicht. Ein Ton der Frequenz 1 kHz hat eine Periodendauer von 0,001 s. Es gibt folgenden mathematischen Zusammenhang zwischen Wellenlänge, Schallgeschwindigkeit und Frequenz beziehungsweise Periodendauer.¹⁷

Formel 2 Wellenlänge λ

$$\lambda = \frac{c}{f}$$

¹⁵ Dickreiter, Michael: Handbuch der Tonstudiotechnik, Bd. 1, Berlin: De Gruyter 2014.

¹⁶ WERSÉNYI, György: Grundlagen der Fahrzeugakustik, in: UNIVERSITAS-GYÖR NONPROFIT Kft., 27.08.2014 [online] <http://vip.tilb.sze.hu/~wersenyi/HTML/GRUNDLAGEN%20DER%20FAHRZEUGAKUSTIK.htm>, [04.02.2022]

¹⁷ Weinzierl 2008: S.23-24.

Zwei identische Schallwellen können eine räumliche oder zeitliche Phasenverschiebung aufweisen. Dann berechnet man die relative Phasenlage oder die Phasenverschiebung $\Delta\phi$. Sie beschreibt den Abstand von zwei gleichen Druckzuständen umgerechnet auf einen Winkel zwischen 0° und 360° . Wird die Phase um 180° verschoben, wird sie als gedrehte Phase oder die Schallwelle als gegenphasig bezeichnet. Räumlicher und zeitlicher Phasenversatz berechnen sich unterschiedlich.¹⁸

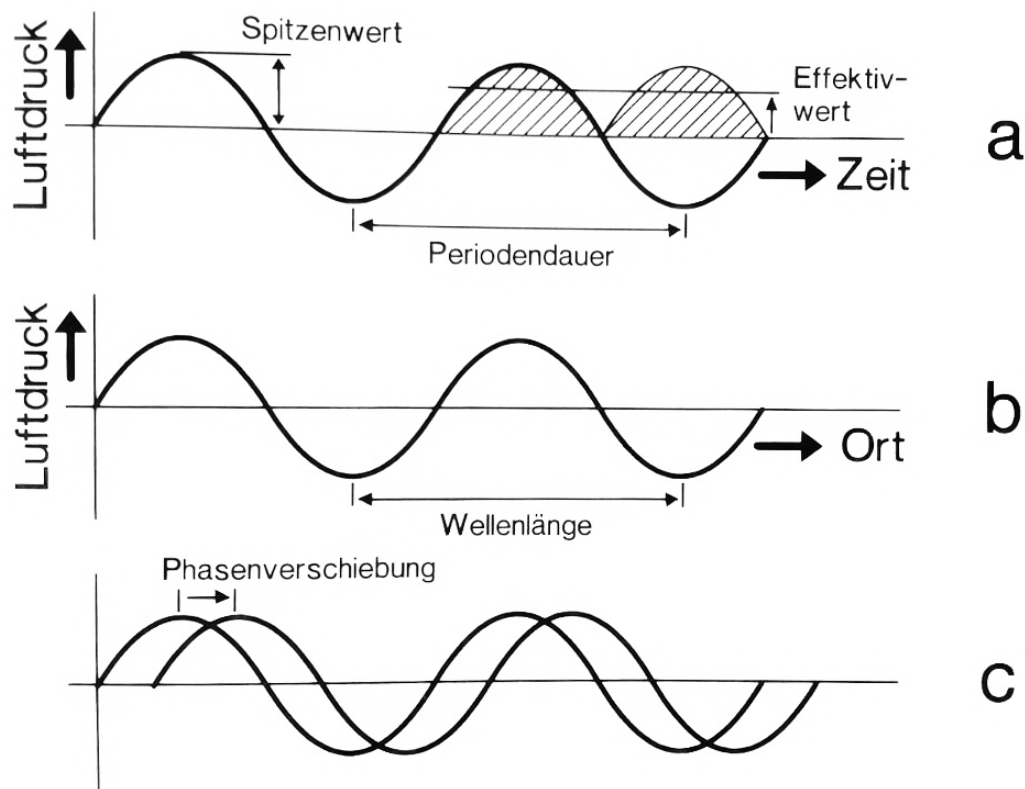


Abbildung 3 Verlauf des Schalldrucks einer Schallwelle im zeitlichen (a) und räumlichen Verlauf (b). Phasenverschiebung zwischen zwei Signalen (c)¹⁹

Die Energie eines Signals wird durch den Effektivwert angegeben. Dieser ist die zeitliche Mittelung eines gleichgerichteten Signals. Somit ist er kleiner als der Maximalwert eines Signals. Normalerweise wird der Schalldruck und die Schallschnelle über einen Effektivwert angegeben. Mitunter wird der Effektivwert in RMS (root mean square) angegeben, was sich auf die Berechnung durch die Wurzel aus dem Mittelwert des Quadrats bezieht.²⁰

¹⁸ Görne 2002: S.15.

¹⁹ Görne 2002: S.15.

²⁰ Warstat, Michael/ Görne, Thomas: Studioteknik Hintergrund und Praxiswissen, 5. Auflage, Aachen: Elektor 2002.

Wenn der Schalldruck in Abhängigkeit mit der Zeit dargestellt wird, erhält man den Zeitverlauf eines Schallsignals. Bei reinen Sinusschwingungen kann man so direkt physikalische Größen, wie die Wellenlänge oder Frequenz, ablesen. Natürliche Klänge, darunter auch die menschliche Stimme, bestehen allerdings aus einer Überlagerung von reinen Schwingungen. So lassen sich keine eindeutigen Größen aus dem Zeitverlauf eines Sprachsignals ablesen. Aus diesem Grund werden Sprachsignale für eine Frequenzanalyse in einem Spektrum dargestellt. Diese Darstellungsform zeigt über einen Zeitabschnitt die Zusammensetzung des Frequenzspektrums für ein Audiosignal. Die Intensität der im Signal vertretenen Frequenzen lassen sich so ablesen. Frequenzbereiche von Signalen können mit Echtzeit-Spektralanalysatoren oder Spectrum Analyzer analysiert werden. Die tiefsten ausgeprägten Frequenzen in einem Spektrum sind die Grundfrequenzen oder -schwingungen. Andere regelmäßige Bestandteile eines Audiosignals sind die Harmonischen oder Obertöne. Die Grundfrequenz bestimmt dabei die subjektiv wahrgenommene Tonhöhe, während die Obertöne die Klangfarbe beeinflussen. Demnach können klangliche Unterschiede mit der unterschiedlichen Frequenzzusammensetzung im Spektrum ermittelt werden. Folglich lässt sich der Unterschied zwischen einem Ton und einem Geräusch ohne tonalen Anteil dadurch erkennen. Je nach Wellenform der Schwingungen sind die Obertöne bestimmte Vielfache der Grundfrequenz. Vielfache der Grundfrequenz erkennt man bei Tönen eindeutiger im Spektrum, während bei tonlosen Geräuschen das Spektrum kontinuierlicher ist.²¹

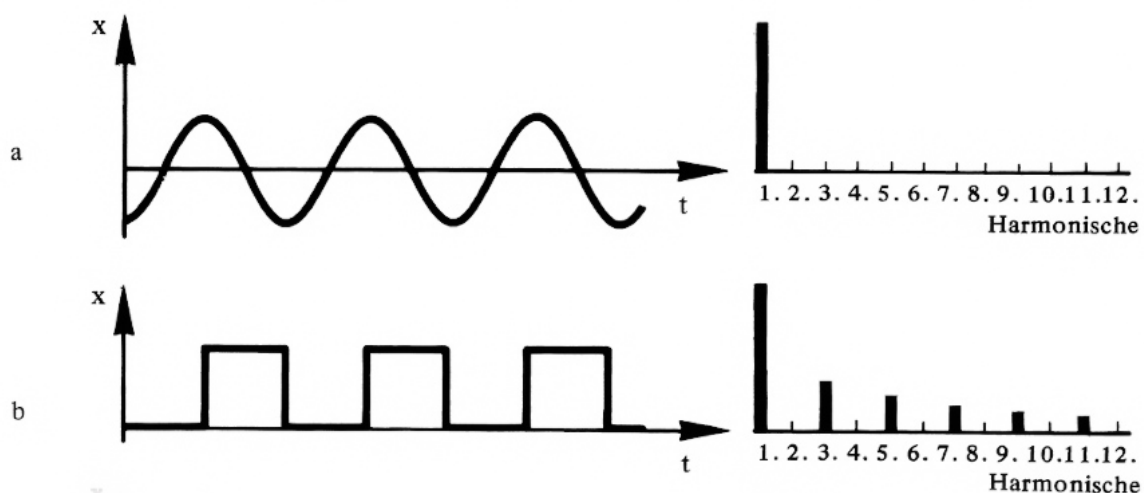


Abbildung 4 Schwingungsverläufe einer Sinuswelle (a) und eines Rechtecksignals (b) mit den dazugehörigen Amplitudenspektren bis zur 12. Harmonischen ²²

²¹ Dickreiter 2014: S.7-9.

²² Dickreiter 2014: S.7.

2.1 Pegel

Der aufzunehmende Schalldruck im Tonstudio bewegt sich zwischen 0,00002 Pa und 200 Pa. In der Audiotechnik wird deshalb statt dem absoluten Wert des Schalldrucks der Schalldruckpegel L_p angegeben. Er ist ein logarithmisches Verhältnismaß und wird in dB (Dezibel) erfasst. Dezibel ist eine Zehntel der Pseudoeinheit Bel des amerikanischen Erfinders Alexander Graham Bell. Der Pegel steht in Zusammenhang mit dem Bezugsschalldruck p_0 . Dieser entspricht dem geradeso noch hörbaren Schalldruck (Hörschwelle) und liegt bei $2 \cdot 10^{-5}$ Pa. Der Schalldruckpegel L_p ist das 20-fach logarithmierte Verhältnis aus dem Schalldruck p zum Bezugsschalldruck p_0 .²³

Formel 3 Schalldruckpegel L_p

$$L_p = 20 \cdot \log\left(\frac{p}{p_0}\right) \text{ dB}$$

Somit gleicht der Schalldruckpegel des Bezugsschalldrucks 0 dB. Der Schalldruckpegel von 200 Pa entspricht 140 dB. Eine Verdopplung des Schalldruckpegels liegt bei +6 dB vor. Die subjektive Lautstärkewahrnehmung verdoppelt sich bei +10 dB. Der Schalldruckpegel in der Audiotechnik wird als SPL (Sound Pressure Level) bezeichnet, da Pegel auch für andere physikalische Größen herangezogen werden. Eine positive Pegelangabe drückt eine Verstärkung aus, während ein negativer Pegel eine Dämpfung darstellt. Ein Minuszeichen wird dem Faktor 20 vorangestellt oder der Kehrbruch aus Schalldruck und Bezugsschalldruck gebildet, um eine Dämpfung zu berechnen.²⁴

Neben dem Schalldruckpegel ist der Spannungspegel L_u im Tonstudio präsent. Er berechnet sich wie der Schalldruckpegel, allerdings werden verschiedene Bezugsspannungen U_0 verwendet. Wenn $U = U_0$ ist, dann liegt der Bezugspegel 0 dBu vor. U_0 ist die effektive Betriebsspannung und setzt an einem Widerstand R von 600 Ohm die Leistung $P = 1$ mW um. Allerdings wird bei der Messung des Spannungspegels kein Widerstand angeschlossen, um einen absoluten Spannungspegel messen zu können. Die Messsituation ist „unloaded“. Es handelt sich um eine unbelastete Situation. Daher wird der Buchstabe u verwendet. Gebräuchlich ist im europäischen Raum die Bezugsspannung 0,775 mV, welche mit dBu oder dBv gekennzeichnet ist. In Japan und den USA liegt sie bei 1 V und wird als dBV bezeichnet. Eine Spannung, die in dBV angegeben ist, ist immer 2,2 dB kleiner als dieselbe Spannung in dBv. Weil ohne die Angabe des an dB nachgestellten Buchstaben keine Rückschlüsse auf die Angabe gezogen werden können, ist diese maßgebend.²⁵ Pegelabweichungen werden in Bezug auf einen willkürlichen Wert angegeben, wie zum Beispiel dem Pegel bei 1 kHz. Unabhängig von der tatsächlichen Spannung wird der Pegel beim Bezugspunkt auf 0 dB gesetzt. Abweichungen werden dann nach obiger Gleichung berechnet.

²³ Smyrek 2012: S.21-22.

²⁴ Weinzierl 2008: S.28-31.

²⁵ Stotz, Dieter: Computergestützte Audio- und Videotechnik. Multimedialechnik in der Anwendung, 3. Auflage, Berlin: Springer 2019.

3. Sprachsignal und Spracherzeugung

Ein grundlegendes Verständnis für die Erzeugung und die auditiven Eigenschaften von menschlicher Sprache begründet in vielen Fällen die Auswahl und den Umgang von Aufnahmetechnik. Deshalb wird in diesem Kapitel auf die menschliche Sprache als Schallereignis und akustisches Sprachsignal eingegangen. Der Mechanismus der Spracherzeugung, sowie eine Frequenzanalyse von menschlicher Sprache sollen die für Sprachaufnahmen relevanten Merkmale der Sprache aufzeigen. Unterschiede zwischen Sprechern und daraus abgeleitete Folgen für Sprachaufnahme schließen das Kapitel ab.

3.1 Akustisches Sprachsignal

Das akustische Sprachsignal ist eine vom menschlichen Sprachapparat abgestrahlte Druckmodulation der Luft. Während im Alltag die Schallwellen den Hörer über das Ohr erreichen, wird das Sprachsignal bei Sprachaufnahmen von einem Mikrofon aufgenommen. Es kann durch eine analoge Echtzeitbeeinflussung bearbeitet werden. Das Sprachsignal steht in Wechselwirkung mit der Raumakustik. Im Fokus dieses Kapitels steht jedoch das akustische Sprachsignal, das aus dem menschlichen Mechanismus der Spracherzeugung resultiert. Die beim Sprechen erzeugten Laute haben Auswirkungen auf die Wahl der Aufnahmetechnik. Je nach individuellem Klang der Laute werden bestimmte Mikrofone und eine Mikrofonposition gewählt oder Maßnahmen gegen Störschall ergriffen, um ein ausgewogenes Klangbild in der Sprachaufnahme hervorzubringen.²⁶ Daher sind praktische Vorgehensweisen bei der Mikrofonierung auf Eigenschaften menschlicher Spracherzeugung zurückzuführen.

Als wesentliches Mittel der Kommunikation, hat sich die menschliche Sprache durch die Evolution weiterentwickelt. Das akustische Signal wurde für die zwischenmenschliche Kommunikation optimiert. Viele Menschen streben im Zuge der technischen Entwicklung danach, Strategien zur Sprachsynthese oder Worterkennung zu entwickeln. Allerdings sind einfache Vorgehensweisen in diesem Bereich nahezu ausgeschlossen. Heutzutage existiert bessere elektronische Messtechnik als zu Beginn der technischen Sprachforschungen Anfang 60er-Jahre. Doch das komplexe Zusammenspiel akustischer Signale untereinander stellt immer noch eine Herausforderung dar. Zahlreiche Phänomene erschweren einen unproblematischen Umgang mit akustischen Sprachsignalen. So gibt es keine eindeutig definierte Trennung zwischen Worten oder Lauten. Aus der Darstellung der Amplitude eines Sprachsignals über die Zeit ist beispielsweise nicht für alle Laute erkennbar, zu welchem Zeitpunkt sie erklingen. Außerdem ist die Lautwahrnehmung menschlicher Sprache abhängig von benachbarten Lauten und hängt von weiteren akustischen Gegebenheiten ab.²⁷

²⁶ Götze, 05.02.2022, siehe Anhang 1.

²⁷ Terhardt, Ernst: Akustische Kommunikation. Grundlagen mit Hörbeispielen, Berlin/ Heidelberg: Springer 1998.

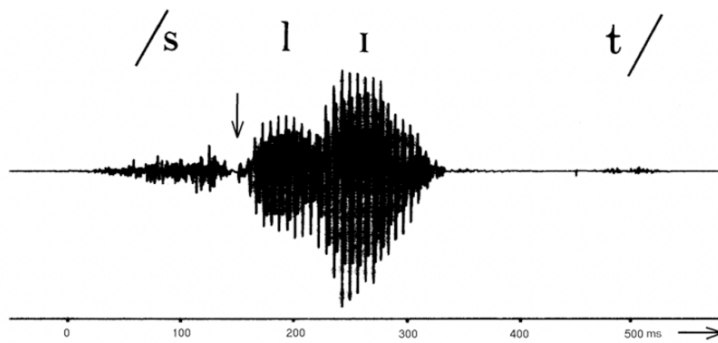


Abbildung 5 Wellenform des Wortes Split. Die markierte Stelle wird als /p/ wahrgenommen²⁸

3.2 Mechanismus der Spracherzeugung

Der Mensch verfügt über einen vielseitigen Mechanismus zur Spracherzeugung. Beim Sprechen wird mechanische Energie aufgebracht und in Schallenergie umgewandelt. Verschiedene Elemente des Vokaltraktes vom Kehlkopf bis zu den Lippen tragen zu einer Vielfalt an Lauten bei. Bei normaler Sprache strömt Luft aus der Lunge. Durch Verengungen im Vokaltrakt kommt es zu stimmloser und stimmhafter Anregung. Resonatoren, die sich im Vokaltrakt befinden, beeinflussen den Luftstrom aus der Lunge so, dass anschließend ein Klang mit einem bestimmten Spektrum von den Lippen abgestrahlt wird. Beim Sprechen äußert der Mensch zahlreiche Laute mit flüchtigem Charakter. Alle Organe zur Spracherzeugung befinden sich in schneller und fast ständiger Bewegung.

Der menschliche Vokaltrakt erzeugt Vokale und Konsonanten. Wird ein Vokal oder Umlaut ausgesprochen, passiert der Luftstrom aus der Lunge einen relativ offenen Vokaltrakt. Wenn Konsonanten artikuliert werden, sorgt ein relativ geschlossener Vokaltrakt dafür, dass der Luftstrom kurzfristig unterbrochen wird. Das ist bei den Verschlusslauten (Plosive) /b/, /p/, /t/ und /k/ der Fall. Bei den Konsonanten /r/, /l/ und /w/ führen Verengungen im Vokaltrakt zu einer kleinen Amplitude des abgestrahlten Sprachsignals. Durch eine ausreichende Verengung entstehen Turbulenzen, wie beispielsweise bei den Reibelauten /f/, /s/ und /z/.²⁹ Die so entstehenden Zisch- oder S-Laute müssen bei der Mikrofonierung miteinbezogen werden.

Beim Erzeugen eines Sprachsignals passiert der Luftstrom zunächst Muskelfalten im Kehlkopf. Die auch als Stimmbänder bezeichneten Falten liegen dicht beieinander. Je nach Spannung und Anregung der Stimmbänder werden unterschiedliche Laute erzeugt. Werden sie durch Luft zum Schwingen angeregt, erzeugen sie einen stimmhaften Laut. Die resultierende Frequenz ist von der Muskelspannung der Stimmbänder abhängig. Für Männer liegt der Mittelwert der erzeugten Frequenzen bei ungefähr 120 Hz, während der gemittelte Wert für Frauen bei etwa 240 Hz einzuordnen ist.³⁰

²⁸ Bernstein, Herbert: Elektroakustik. Mikrofone, Klangstufen, Verstärker, Filterschaltungen und Lautsprecher, 2. Auflage, Wiesbaden: Springer 2019.

²⁹ Bernstein 2019: S.49.

³⁰ Ernst 1998: S.30-33.

Werden die Stimmbänder angeregt, wenn sie weiter auseinander liegen, folgt ein stimmloser Laut. Beim Flüstern sind ausschließlich stimmlose Laute zu vernehmen. Nachdem die Luft die Stimmbänder transitiert hat, beeinflussen Artikulatoren den Klang des stimmhaften oder stimmlosen Lautes. Kiefer- und Zungenmuskeln bewegen die Rachenhöhle, das Zäpfchen, das Gaumensegel, den Gaumen und den Zahndamm (Alveolen) im Verhältnis zu den übrigen Elementen des Vokaltraktes. Die dadurch erzeugte Wirkung auf den Luftstrom und die durch den Vokaltrakt erzeugten Resonanzen bestimmen das Klangbild des sich ergebenden Lautes. Abhängig von der Position der Verengungsstellen im Mechanismus zur Spracherzeugung können unterschiedliche Artikulationsarten angenommen werden.³¹

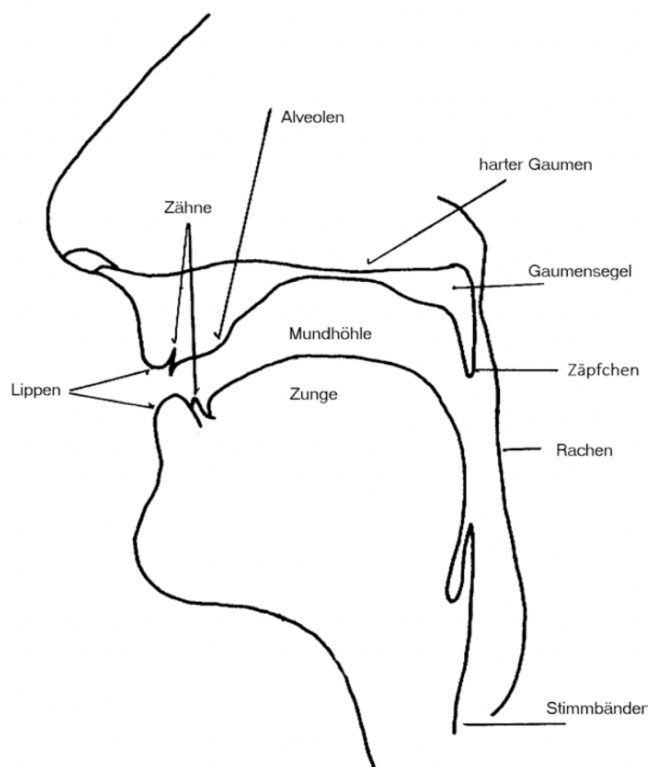


Abbildung 6 Mechanismus zur Spracherzeugung beim Menschen ³²

In der deutschen Sprache zählen unter anderem Reiblaute zu den Artikulationsarten. Entweicht die Luft auf einer oder mehreren Seiten der Zunge entsteht ein Seitenlaut (/l/, /r/). Bei den Verschlusslauten /m/ und /n/ ist der Vokaltrakt und der Nasenraum geschlossen. Wird kurzfristig viel Druck durch einen geöffneten Nasenraum bei Verschlusslauten frei, werden Plosive (/p/, /b/, /t/, /d/, /k/, /g/) hervorgebracht. Gegen den kurzfristig starken Luftimpuls müssen Maßnahmen gegen Störschall ergriffen werden. Ein gleitender Übergang von einem Verschlusslaut in einen Reiblaut definiert Affrikaten (/ts/ wie in „Katze“). Bei Approximanten führt eine annähernde Artikulation zu einer Mischform aus Konsonanten und Vokalen.³³

³¹ Ernst 1998: S.30-33.

³² Bernstein 2019: S.50.

³³ Bernstein 2019: S.51.

Neben der Stellung der Elemente im Vokaltrakt, beeinflusst die relative Bewegung oder Position eines Artikulators Laute bei der Spracherzeugung. Beispielsweise kann die Zunge mehr als zwei Positionen für eine gleiche Stellung des restlichen Sprachmechanismus einnehmen. Zusätzlich verfügt die Lippe über weitere Funktionen wie zum Beispiel die Verlängerung des Vokaltraktes. So lassen sich Vokale mit unterschiedlicher Länge bilden (/a/ in „schlaff“ im Vergleich zu anders klingenden /a/ in „Wagen“).³⁴

Bei normaler und schnell gesprochener Sprache besteht eine, durch verschiedene Phänomene hervorgerufene, physikalische Trägheit. Dadurch werden artikulierte Laute nicht vollkommen isoliert voneinander artikuliert. Stattdessen erreichen die Artikulatoren nur näherungsweise ideale Positionen. Folglich überlappen sich aufeinanderfolgende Laute, was als Koartikulation bezeichnet wird. Darüber hinaus klingen Vokale auf Grund der Vokalreduktion in vielen Fällen akustisch ähnlicher als vermutet. Das liegt daran, dass auch der Vokaltrakt beim alltäglichen Sprechen keine idealen Positionen einnimmt. Beide Phänomene beeinträchtigen die Sprachverständlichkeit.³⁵

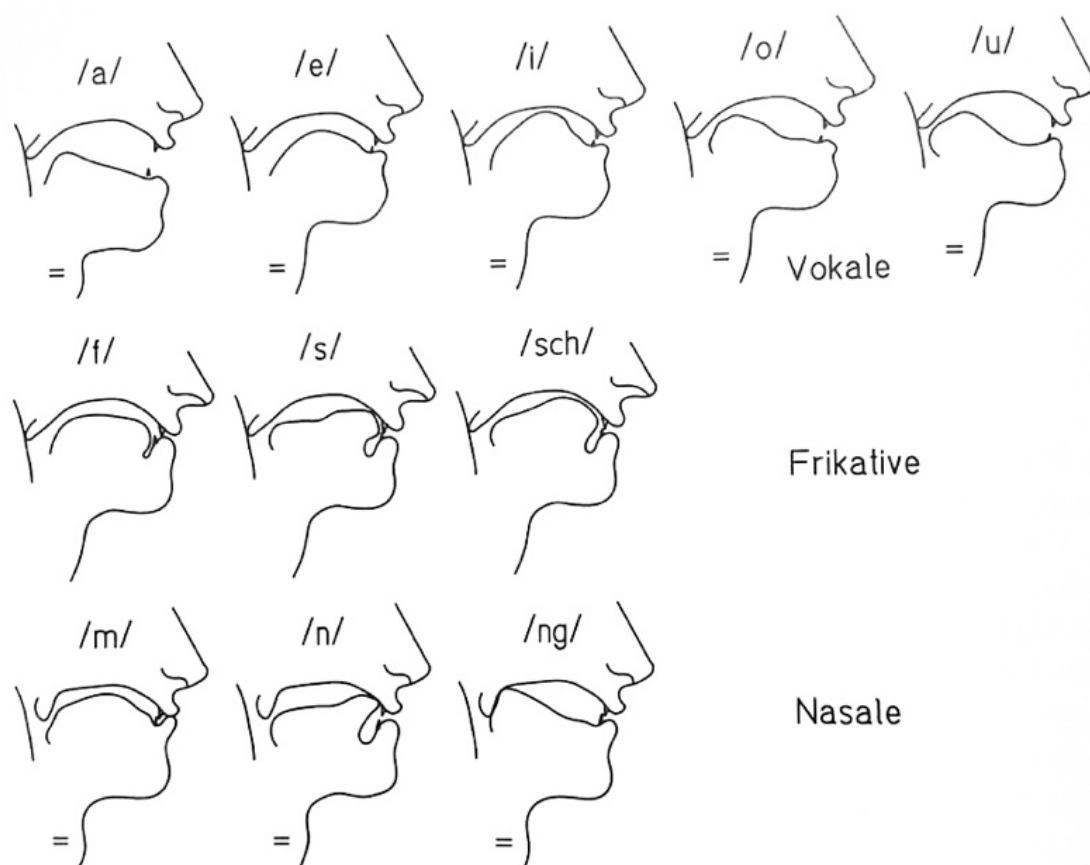


Abbildung 7 Artikulation von Lauten als schematische Darstellung³⁶

³⁴ Bernstein 2019: S.51.

³⁵ Ernst 1998: S.183-184.

³⁶ Ernst 1998: S.184.

In der Sprachforschung und der Phonetik zieht man die Mittel zur Lautzeugung in Systeme zur Sprachsynthese oder -analysierung mit ein. Heutzutage liegt der Fokus verstärkt auf der parametrischen Darstellung des akustischen Sprachsignals. Im Bereich Hörbuchproduktion ist ein Verständnis für die Spracherzeugung elementar. Das fachliche Vokabular ermöglicht eine effektive Kommunikation mit Sprechern. Eine profunde Auffassung des aufzunehmenden Schallereignisses ist für den gesamten Vorgang einer Hörbuchproduktion von Vorteil.

3.2.1 Quelle-Filter-Modell

Die physikalischen Wirkungen bei der menschlichen Spracherzeugung können vereinfacht durch ein Quelle-Filter-Modell (source-filter model) simuliert werden. Das technische Modell liefert mittels linearer Filter eine hinreichend gute Annäherung für wichtige Sprachmerkmale. Unabhängig von den Filtern liefert eine Quelle eine Anregungsfunktion, die periodisch ist oder ein Rauschen aufweist. Im Regelfall werden rein periodische Signale oder Rauschsignale für die Lauterzeugung generiert. Für Laute mit stimmhafter und periodischer Anregung (/z/, /v/) wird ein periodisches Anregungssignal und danach stimmlose Frikative (/s/, /f/) mit einem Rauschsignal erzeugt. Quelle-Filter-Modell erzeugte Mischlaute weisen allerdings eine schlechtere Qualität als andere simulierte Laute auf. Die Periode der Anregungsfunktion wird verändert, um das Wechseln der von den Stimmbändern erzeugten Grundfrequenzen nachzuahmen. Dieses Signal enthält bereits viele Spektralkomponenten.

Filter manipulieren den Klang des von der Anregungsfunktion hervorgebrachten Signales. Vergleichbar ist ihre Wirkung mit den Artikulatoren im Vokaltrakt, die den Ton der Stimmbänder beeinflussen. Verschiedene Filterdurchlasskurven führen zu unterschiedlichen Lauten. Die Durchlasskurven unterscheiden sich für Vokale und Konsonanten bedeutend. Wie die Filter im menschlichen Vokaltrakt, ändern die simulierten Filter die Tonhöhe des Lautes nicht.³⁷

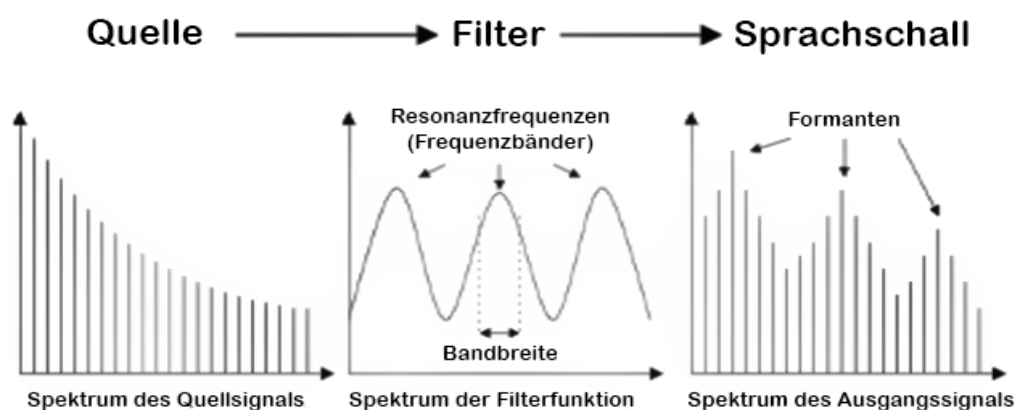


Abbildung 8 Quelle-Filter-Modell für Spracherzeugung³⁸

³⁷ Bernstein 2019: S.64-69.

³⁸ Trommer, Jochen: Resonanz. Das Quelle-Filter-Modell der Vokalproduktion, in: Universität Leipzig Institut für Linguistik, 2007 [online] <https://www.uni-leipzig.de/>, [04.02.2022]

3.2.2 Formant-Syntesizer

Die Funktionsweise eines Formant-Syntesizers baut auf das Grundprinzip des Quelle-Filter-Modells auf. Entwickler des Formant-Syntesizers stützen sich auf die These, dass Formanten zur Synthese von verständlicher Sprache verwendet werden können. Ein Formant bezeichnet einen Oberton oder zusammenhängenden Frequenzbereich von Obertönen, die die Klangfarbe eines Lautes beeinflussen.³⁹ Folglich sind stimmlose Laute keine Formanten. Der Begriff kommt vom lateinischen „formare“, was formen bedeutet. Das Quelle-Filter-Modell soll ein Sprachsignal simulieren, indem für Laute charakteristische Resonanzfrequenzen angegeben werden.

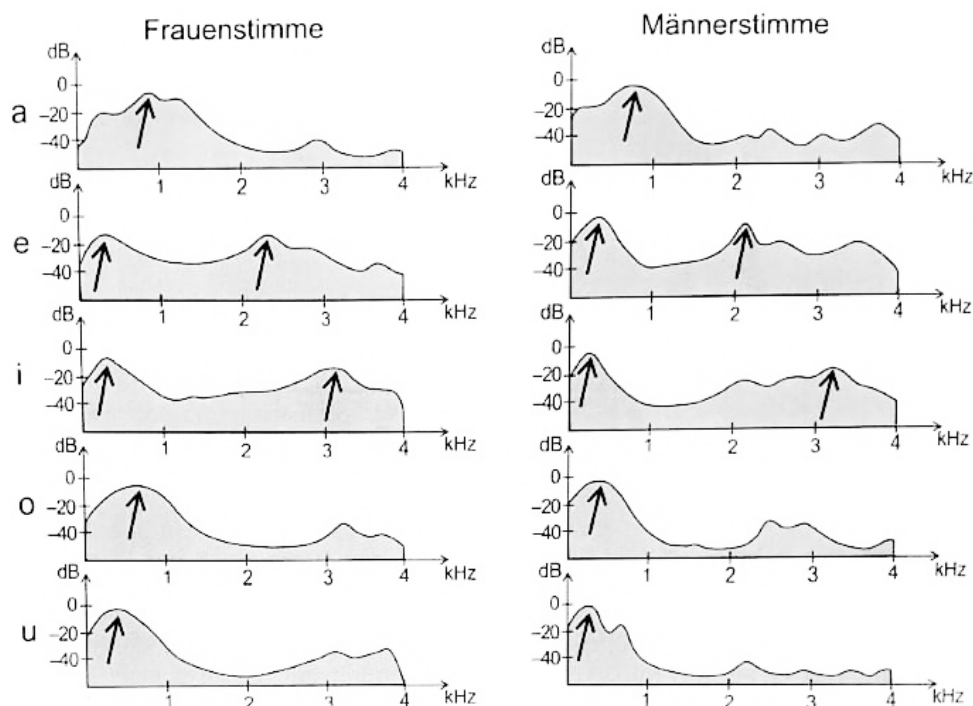


Abbildung 9 Frequenzspektren (Teiltonspektren) von Vokalen und ihre Resonanzfrequenzen ⁴⁰

Durch eine geringe Anzahl von in der Regel fünf Resonanzen erzeugt ein Formant-Syntesizer ein verständliches Sprachsignal. Je mehr Filter miteinbezogen werden, desto detaillierter wird das Sprachsignal simuliert. Für jeden Formanten ist ein separater Filter in analogen Synthesizern notwendig. Die Filter manipulieren die Frequenz, Amplitude und Bandbreite des Quellsignals. In der Praxis werden die Filter in Reihe geschaltet, da dadurch eine bessere Annäherung an die realen Filtercharakteristiken erreicht wird. Das liegt mitunter daran, dass bei der Parallelschaltung von Resonanzfiltern die Amplituden jedes Formanten separat eingestellt werden muss. Bei der Reihenschaltung können Amplitude, Frequenz und Bandbreite von jedem Filter in Abhängigkeit zueinander beeinflusst werden.⁴¹

³⁹ Pawera, Norbert: Mikrofonpraxis. Tipps und Tricks für Bühne und Studio; Technik, Akustik und Aufnahmepraxis für Instrumente und Gesang, 4. Auflage, München: Franzis 2003.

⁴⁰ Smyrek, Volker: Tontechnik. Für Veranstaltungstechniker in Ausbildung und Praxis, 2. Auflage, Stuttgart: Hirzel 2012.

⁴¹ Bernstein 2019: S.64-69.

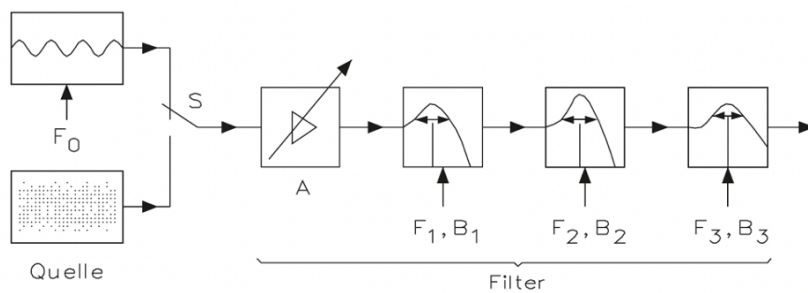


Abbildung 10 Blockschaltbild von Sprachsynthesizer mit drei Formanten in Reihenschaltung⁴²

3.3 Frequenzanalyse des Sprachsignals

Mit einer Frequenzanalyse können akustische Sprachsignale untersucht werden. Die Analyse gibt Rückschlüsse auf die Wechselwirkung zwischen dem Mechanismus der Spracherzeugung und dem resultierenden Schallereignis. Mit einem Frequenz-Zeit-Spektrogramm werden Frequenzanteile von Sprache dargestellt. Die Energiespitze, die durch die entweichende Luft vom Menschen abgestrahlt wird, wird im Frequenz-Zeit-Spektrogramm veranschaulicht. Dadurch können die durch den Vokaltrakt hervorgebrachten wesentlichen Resonanzen für Formanten identifiziert werden. Die Analyse erschließt den Frequenzbereich, der für einen Formanten charakteristisch ist (Formantbereich). Anhand der Resonanzfrequenz, Amplitude und Bandbreite kann ein bestimmter Formant einem Zeitabschnitt des Spektrogramms zugeordnet werden.⁴³ Formanten bestimmen den Klangcharakter einer Stimme und sind bei jedem Sprecher individuell.

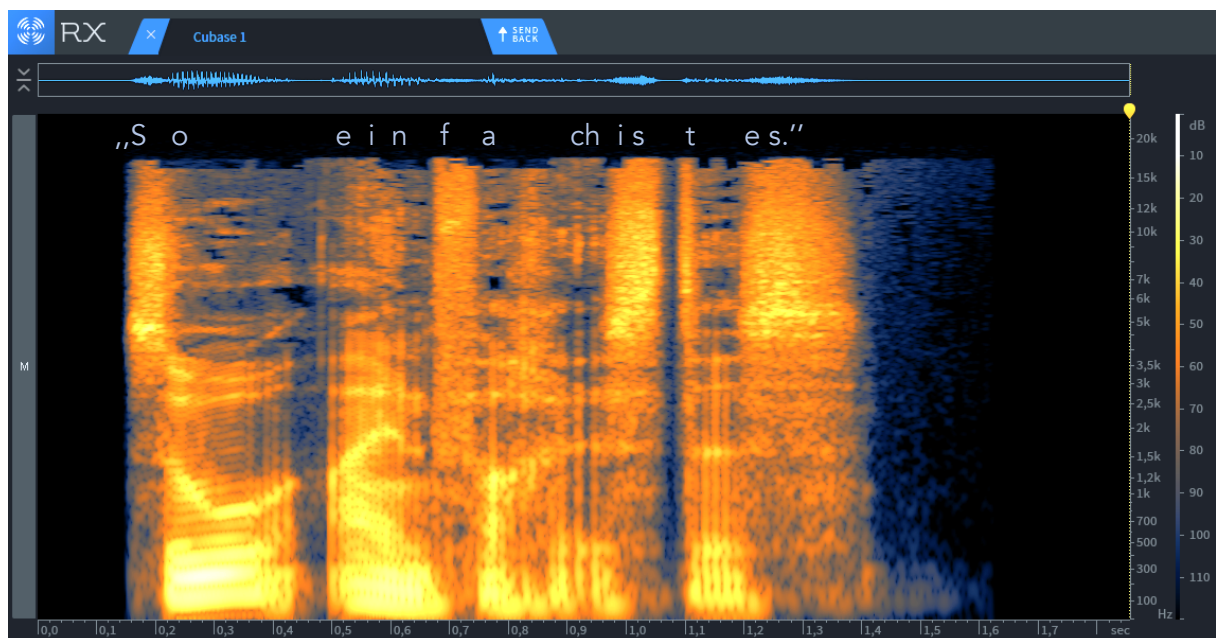


Abbildung 11 Screenshot aus iZotope RX Connect: Frequenz-Zeit-Diagramm des Satzes „So einfach ist es.“

⁴² Bernstein 2019: S.68.

⁴³ Ernst 1998: S.

Aus der Frequenzanalyse resultiert unter anderem, dass einleitende Konsonanten die folgenden Formanten nicht beeinflussen. Beispielsweise weist /di/ einen höheren Formanten als /du/ auf. Je weiter vorne sich eine Verengung im Vokaltrakt befindet, desto höher erklingt die Resonanzfrequenz. Das liegt daran, dass der Resonanzraum des Vokaltraktes im vorderen Abschnitt kleiner wird. Aus der Frequenzanalyse mit einem Frequenz-Zeit-Diagramm geht auch hervor, dass auf Plosivlaute eine ungefähr 50 ms lange relative Ruhe folgt.⁴⁴

In der Aufnahmepraxis wird in der Regel keine Frequenzanalyse für den Sprecher durchgeführt. Dennoch belegen Zeit-Frequenz-Spektrogramme, dass Laute ungefähr in bestimmte Frequenzbereiche eingeordnet werden können. Diese Tatsache ist für Sprachaufnahmen insofern interessant, da sich dieselben Laute von verschiedenen Sprechern in andere Frequenzbereiche einordnen. Auch die Frequenzbereiche von stimmlosen Lauten wie S- und Zischlaute können variieren.⁴⁵

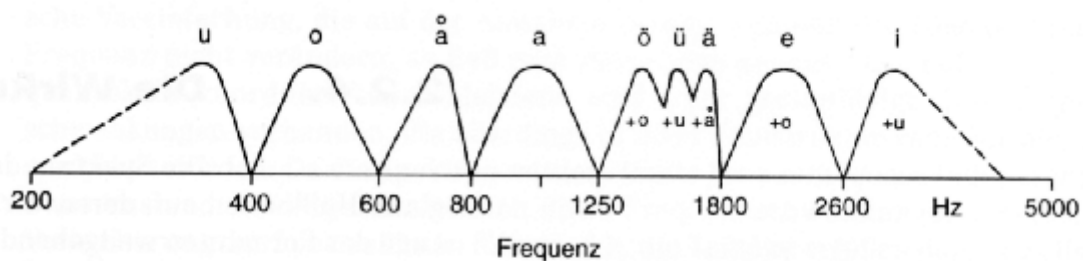


Abbildung 12 Durchschnittliche Frequenzbereiche für unterschiedliche Vokale⁴⁶

3.4 Unterschiede zwischen Sprechern

Aus verschiedenen Gründen hört sich die Sprache diverser Sprecher unterschiedlich an. Individuelle Stimmbandgrundfrequenzen und die Größe des Vokaltraktes bestimmen den Klang einer Stimme. Mit einem größeren Vokaltrakt gehen größere Resonanz-Hohlräume einher, wodurch tiefere Frequenzen erzeugt werden. Anatomisch begründet ist der Kontrast zwischen Kinder- und Erwachsenenstimmen oder Männer- und Frauenstimmen sehr ausgeprägt. Ferner führt die Vokaltraktform und Unterschiede in der Zusammenwirkung einzelner Elemente des Sprachmechanismus zu verschiedenen akustischen Ergebnissen.⁴⁷

⁴⁴ Bernstein 2019: S.52-56.

⁴⁵ Götze, 05.02.2022, siehe Anhang 1.

⁴⁶ HfMDK Frankfurt: Töne, Klänge und Klangfarben. Kapitel 1. Schwingungen und Wellen, in: Akustische Grundlagen, 2022 [online] https://www.selma.hfmdk-frankfurt.de/akustische-grundlagen/abb/tut_formant.png, [07.02.2022]

⁴⁷ Dickreiter 2014: S.69-70.

In einem Abstand von etwa 30 cm liegt der durchschnittliche Schalldruckpegel eines Sprechers bei 64 dB. Wird der Abstand um die Hälfte verringert, erhöht sich der Schalldruckpegel auf jeweils um 6 dB.⁴⁸ Es ist sinnvoll sich auf statistisch ermittelte Durchschnittswerte der Sprache, anstatt auf einzelne Laute zu beziehen. Auch der durch Sprache hervorgerufene Schallpegel (Sprachschallpegel) variiert bei Sprechern. Ausgebildete Sprecher bemühen sich um einen zeitbezogen gleichmäßigen Lautstärkeverlauf. Ursache dafür ist, dass laute Sprache obertonreicher ist und somit anders klingt.⁴⁹

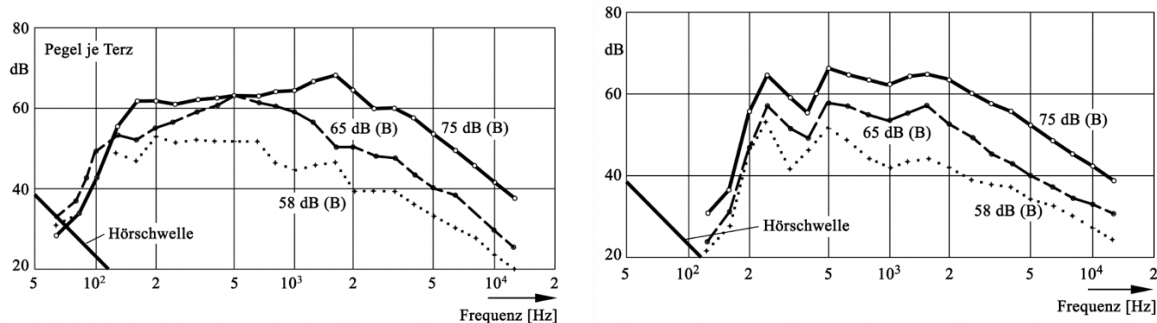


Abbildung 13 Durchschnittliche Spektren deutscher Sprecher für unterschiedliche Lautstärken. Männliche Sprecher (links) und weibliche Sprecher (rechts)⁵⁰

Besonders deutlich ist der hörbare Unterschied zwischen qualifizierten Sprechern und Laien. Deutliches und betonendes Sprechen über einen längeren Zeitraum erfordert einen trainierten Sprachapparat. Laien unterschätzen häufig den Anspruch an Hörbuchsprecher. Ihnen sind Spracheigenschaften in der Regel nicht bewusst und beachten deshalb keine professionellen Sprechregeln. Geschulte Sprecher meiden unerwünschte Nebengeräusche wie überdurchschnittlich lautes Atmen, sprachlose Mundgeräusche und hörbare Körperbewegungen beim Sprechen. Außerdem sind berufliche Sprecher im Vergleich zum durchschnittlichen Leser im zu lesenden Text voraus. Infolgedessen hat sich ein Sprecher bereits eine passende Betonung ausgedacht, bevor er vorliest.

In der Praxis muss auf die unterschiedlich klingende Sprache eingegangen werden. Je nach klanglichen Besonderheiten wird eine bestimmte Aufnahmetechnik gewählt. Des weiteren richten sich Einstellungen und die Aufstellung der Aufnahmetechnik individuell nach der Sprecherstimme. Ziel dabei ist es die Sprachaufnahme harmonisch zu halten. Beispielsweise können so klangliche Defizite wie überdurchschnittlich laute S-Laute oder stark betonte Plosive vermindert werden.⁵¹

⁴⁸ Dickreiter 2003: S.188.

⁴⁹ Dickreiter 2014: S.69-70.

⁵⁰ Dickreiter 2014: S.69.

⁵¹ Götze, 05.02.2022, siehe Anhang 1.

4. Mikrofonierung

Qualitativ hochwertige Sprachaufnahmen bedürfen einer angemessenen Mikrofonierung. Das Ziel der Mikrofonierung von Sprache ist es, die Schallschwingungen so in elektrische Spannung umzuwandeln, dass der Klang der Stimme positiv beeinflusst wird. Eine signalgetreue Schallwandlung ist ein angestrebter Idealfall, der nicht unbedingt bei Aufnahmen im Bereich Hörbuch erforderlich ist. Mit der Aufnahme einhergehende Sprachverzerrungen sollten einen ausgewogenen Klang der Stimme unterstützen.⁵² Da die Schallwandlung in einem Schritt nicht machbar ist, wird Schall in mechanische Schwingungen und anschließend in ein elektronisches Signal gewandelt. Der erste Schritt findet in der Mikrofonkapsel statt und wird beeinflusst durch die Richtcharakteristik des Mikrofons. Danach findet die elektromechanische Wandlung bei Sprachaufnahmen in der Regel nach der Funktionsweise von Kondensatormikrofonen statt. Darüber hinaus wirkt sich die Position des Mikrofons gewichtig auf das aufgenommene Audiosignal aus. Dabei treten vielerlei Phänomene auf. Aus diesen leiten sich praktische Vorgehensweisen ab, die sich bei Sprachaufnahmen bewährt haben. Dieses Kapitel soll die Besonderheiten bei der Mikrofonierung bei Hörbuchaufnahmen erläutern und die daraus resultierende Mikrofonpraxis schildern.

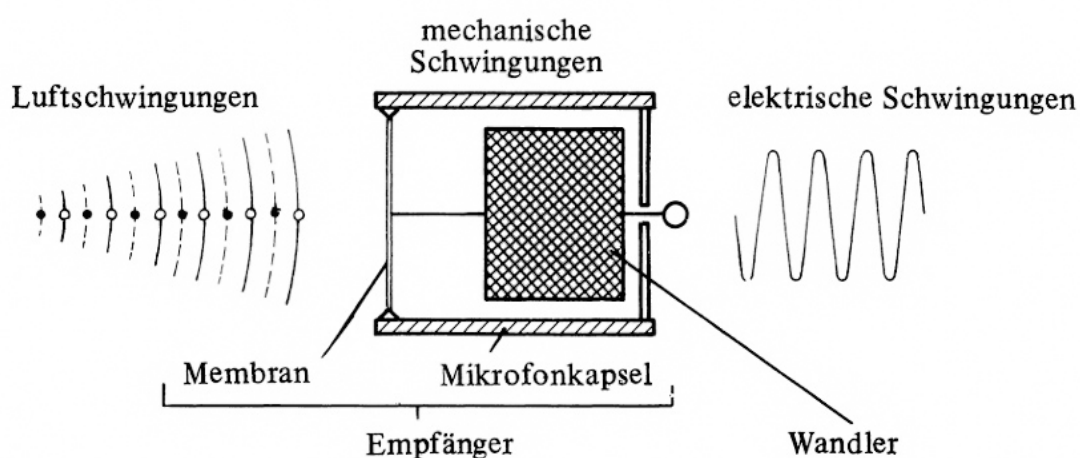


Abbildung 14 Aufbau eines Mikrofons ⁵³

Sprachaufnahmen für Hörbücher erfordern keine mehrkanalige Aufnahmetechnik. Es reicht aus, ein Mikrofon einzusetzen und einkanalig aufzunehmen. Für einen Vergleich mehrerer Mikrofone kann man dennoch mit unterschiedlichen Mikrofonen aufnehmen. Allerdings wird im finalen Hörbuch nur ein Mikrofon zu hören sein. Deshalb hat die Monotechnik als Ursprung der Tonaufnahmetechnik in Bezug auf Sprachaufnahmen für Hörbücher immer noch seine Daseinsberechtigung.⁵⁴

⁵² Götze 04.02.2022, siehe Anhang 1.

⁵³ Dickreiter 2014: S.139.

⁵⁴ Henle, Hubert: das Tonstudio Handbuch, 5. Auflage, München: GC Carstensen, 2001.

4.1 Richtcharakteristik

Die Richtcharakteristik gibt die Richtungsabhängigkeit der Mikrofonempfindlichkeit an. Aus ihr lassen sich bevorzugte Aufnahmerichtungen ableiten. Das Richtverhalten eines Mikrofons wird von der Bauweise der Mikrofonkapsel, der mechanisch-akustischen Konstruktion, bestimmt. Bei Mikrofonen im Tonstudio kann die Kapsel oft gewechselt oder die Richtcharakteristik umgeschaltet werden. Bei jeder Mikrofonkapsel ist die Membranvorderseite grundsätzlich zugänglich für den Schall.

Die einfachste Kapselform ist die Druckempfängerkapsel (Pressure Transducer). Ist die Membranrückseite geschlossen, wodurch der Schall auf der Membranrückseite nicht auftreten kann, handelt es sich um einen Schalldruckempfänger.⁵⁵ Die Kapselmembran wird vom Schalldruck bewegt. Die Membran bewegt sich bei steigendem Luftdruck nach innen und abnehmendem Druck nach außen. Wechselt der Luftdruck, schwingt die Membran. So wird die akustische Schallenergie in mechanische Energie umgewandelt. Schalldruckempfänger sind ungerichtet, da die Membran unabhängig von der Richtung, aus der die Schallwellen die Kapsel erreichen, schwingt. Solange die Wellenlängen des Schalls größer als der Durchmesser der Mikrofonkapsel sind, tritt dieser Effekt auf. Daraus ergibt sich eine gleiche Empfindlichkeit in alle Richtungen. Druckempfänger haben deshalb eine Kugelcharakteristik und wandeln omnidirektional Schallwellen um. Tiefe Frequenzen werden besonders signalgetreu von Mikrofonen mit Druckempfängerkapsel umgewandelt. Eine Druckausgleichsöffnung (Kapillare) legt durch ihre Größe die akustische untere Grenzfrequenz fest. Ferner verhindert sie, dass die Membran einen Ausschlag in Hochdruckgebieten aufweist.⁵⁶ Hohe Frequenzen aus der Einfallrichtung 0° werden von der Membran reflektiert. Der daraus entstehende Druckstau führt zu einer Höhenanhebung um bis zu 10 dB. Aus diesem Grund verändert sich die Richtcharakteristik von Druckempfängern zu höheren Frequenzen im Vergleich zu Druckgradientenempfängern weniger.⁵⁷

Bei Hörbuchaufnahmen werden hauptsächlich Druckgradientenempfänger eingesetzt.⁵⁸ Die Kapsel eines Druckgradientenempfängers erlaubt es, dass Schall beide Seiten der Membran erreichen kann. Weil die Differenz zwischen dem Schall, der auf der Vorder- und Hinterseite der Membran auftritt, die Richtungsempfindlichkeit der Kapsel bestimmt, ist auch der Begriff Druckdifferenzempfänger üblich. Je nach Kapselbauweise und Richtung, aus der der Schall kommt, treffen die Schallwellen mit zeitlichem Versatz zueinander auf beiden Seiten der Membran auf. Dadurch entsteht ein Phasenunterschied. Es ergibt sich ein Druckunterschied, der vom Einfallswinkel der Schallwellen abhängig ist. Aus dieser Kapselbauweise resultiert, dass alle Druckgradientenempfänger über eine richtungsabhängige Schallempfindlichkeit verfügen.

⁵⁵ Smyrek 2012: S.146.

⁵⁶ Dickreiter 2014: S.152-154.

⁵⁷ Henle 2001: S.153.

⁵⁸ Götze 04.02.2022, siehe Anhang 1.

Der Druckunterschied wird durch den Einsatz schallverzögernder Elemente manipuliert. Eine zweite Membran, Schallkanäle, Luft- oder Gazeöffnungen beeinflussen als akustische Laufzeitglieder die Differenz der auf den Membranseiten auftreffenden Schallwellen. Verbaute schalldurchlässige Konstruktionen führen beispielsweise zu einer Nierencharakteristik. Überlagert man eine Acht- mit einer Kugelrichtcharakteristik können auch Druckgradientenempfänger mit Nieren-, Super- und Hypernierenrichtcharakteristik konstruiert werden.⁵⁹

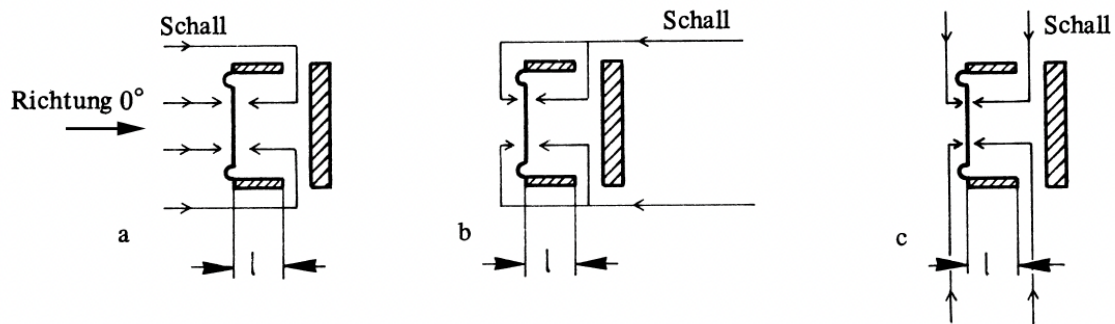


Abbildung 15 Kapselaufbau eines Druckgradientenempfängers mit akustischen Laufzeitgliedern l und Nierencharakteristik. Verschiedene Schalleinfallsrichtungen (a, b, c) ⁶⁰

Eine positive Halbwelle einer Schallwelle lenkt die Membran von Druckgradientenempfängern, je nach Einfallrichtung, nach vorne oder hinten aus. Es gibt hier also eine richtungsabhängige Fallunterscheidung. Verglichen mit einem Druckempfänger, würde sich dessen von einer positiven Schallwelle angeregte Membran nach hinten beziehungsweise ins Kapselinnere bewegen. Dabei beeinflusst die Richtung, aus der der Schall kommt, die Membranbewegung nicht.⁶¹

Anders als bei Druckempfängern ist die Antriebskraft der Gradientenempfänger frequenzabhängig. Mit abnehmender Frequenz verkleinert sich der Druckgradient. Das hat zur Folge, dass sehr tiefe Frequenzen von Druckdifferenzempfängern unsauber wiedergegeben werden. Der Druckgradient steigt für höhere Frequenzen zunächst an. Wenn die mittlere Wegdifferenz zwischen Membranvorder- und rückseite einer halben Wellenlänge oder der Phasendifferenz 180° entspricht. Bei größerer Phasendifferenz wird die Membranauslenkung wieder kleiner, bis sie bei 360° nicht mehr ausgelenkt wird. Danach steigen der Druckgradient und die die Auslenkung erneut an.⁶²

⁵⁹ Henle 2001: S.157.

⁶⁰ Dickreiter 2014: S.158.

⁶¹ Pawera 2003: S.46-47.

⁶² Webers 2007: S.255-259.

Die mittlere Wegdifferenz zwischen Membranvorder- und -rückseite ist so festgelegt, dass sie etwa der halben Wellenlänge der höchsten zu übertragenden Frequenz entspricht. So wird eine Membranauslenkung bewirkt, die mit der Frequenz zunimmt. Das Mikrofon arbeitet abgesehen von seinem oberen Übertragungsbereich im ansteigenden Teil der Frequenzkurve. Oberhalb der Frequenz $f_{\text{Ü}}$ wird das Druckgradientenprinzip unwirksam. An dieser Stelle wird das Mikrofon zu einem Druckempfänger. Es wird eine Dämpfung eingesetzt, deren Wirkung ebenfalls mit der Frequenz zunimmt. Dadurch wird das Ansteigen der Membranauslenkung von tiefen zu hohen Frequenzen korrigiert.⁶³

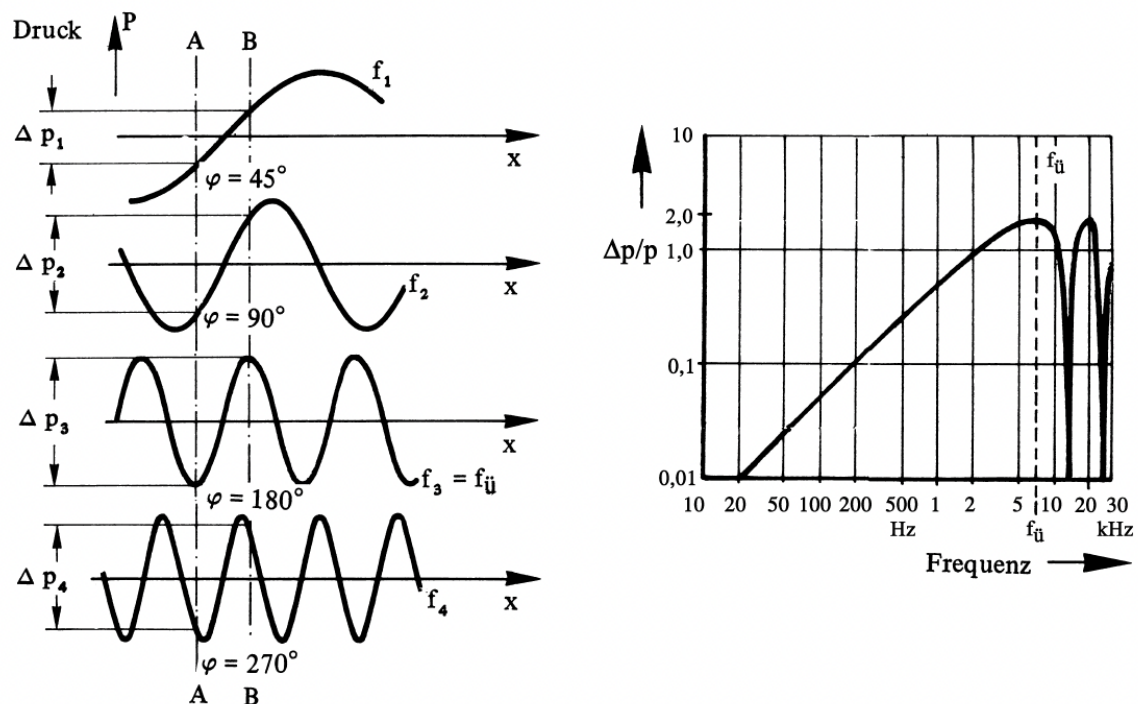


Abbildung 16 Schalldruckdifferenzen bei verschiedenen Frequenzen und Frequenzgang des Druckgradienten D_p zwischen der Membranvorderseite (A) und -rückseite (B)⁶⁴

Eine Massehemmung und Reibungshemmung bewirken eine frequenzabhängige Dämpfung. Bei der Massehemmung entsteht eine der sich bewegenden Masse entgegenwirkende Trägheit. Je schneller sich die Masse bewegt, desto größer ist dieser Widerstand. Die Membranmasse erfährt einen zunehmenden Widerstand, je größer und schneller Bewegungen erfolgen. Verursacht durch die Membranbewegungen strömt Luft durch die Öffnungen der Kapselrückseite, wodurch Reibungshemmung entsteht. Luftmoleküle, die an den Wänden der Öffnungen reiben bringen der Membranbewegung einen weiteren Widerstand entgegen. Die Reibung ist umso stärker, je schneller die Luft strömt. Die Reibungshemmung steigt wie die Massehemmung, mit zunehmender Frequenz.⁶⁵

⁶³ Webers 2007: S.255-259.

⁶⁴ Boré, Gerhart/ Peus, Peter: Mikrophone. Arbeitsweise und Ausführungsbeispiele, 5. Auflage, Berlin: Georg Neumann GmbH 2002.

⁶⁵ Dickreiter 2014: S.154-156.

Je frequenzunabhängiger die Richtcharakteristik eines Mikrofons ist, desto weniger Verfärbungen treten bei Bewegungen während des Sprechens auf. Dreht ein Sprecher seinen Kopf beim Lesen, beeinflusst die Richtcharakteristik die Qualität der Aufnahme. Bei der Kopfdrehung wird der Schall am Mikrofon vorbei abgestrahlt. Nimmt man mit einer breiten Niere auf, fällt der klangliche Unterschied bei Bewegungen im Vergleich zu einer Superniere weniger auf.⁶⁶ Dazu kommt, dass eine stark frequenzabhängige Richtcharakteristik bei Kopfbewegungen die Sprache verfärbt. Die Richtcharakteristik einer Niere neigt dazu, sich zu höheren Frequenzen mehr zu einer Acht und zu tieferen Frequenzen zu einer Kugel zu verzerren.⁶⁷ Wenn tiefe Frequenzen aus einer bestimmten Richtung noch mit gleichem Pegel wie bei direkter Besprechung aufgenommen werden und hohe Frequenzen nicht, klingt die indirekte Besprechung dumpf. Verwendet man dennoch ein stark gerichtetes Mikrofon mit ausgeprägten Abweichungen in der frequenzabhängigen Richtcharakteristik, sollte der Sprecher stets darauf achten den Kopf nicht zu drehen. Außerdem verringert das Lesen von einem Tablet, statt von einer gedruckten Doppelseite, wiederkehrende Kopfdrehungen.

Bei Sprachaufnahmen für Hörbuchproduktionen werden Mikrofone mit Richtwirkung eingesetzt, um Direktschall aus bevorzugten Richtungen aufzunehmen. Störschall soll möglichst nicht aufgenommen werden. Weit über den Hallradius ist die Richtwirkung nicht mehr wirksam. In sehr halligen Umgebungen und bei vielen Störgeräuschen empfiehlt sich daher ein sehr gerichtetes Mikrofon. Bei Sprachaufnahmen im Tonstudio sind keine stark gerichteten Mikrofone notwendig. Je nach Richtcharakteristik tritt ein bestimmtes Dämpfungsverhalten für entsprechende Richtungen auf. Folglich unterscheidet sich die Dämpfung von Mikrofonen mit den Richtcharakteristiken Kugel, Keule, Niere, Superniere, Hyperniere und Acht. Bei Sprachaufnahmen für Hörbücher im Tonstudio werden hauptsächlich Mikrofone mit einer Nierencharakteristik (Cardioid) eingesetzt. Seitlicher Schall wird um 6 dB gedämpft und rückwärtiger Schall wird völlig ausgelöscht. Das hat den Vorteil, dass weniger Raum aufgenommen wird. Die relative Phasenlage ist für den gesamten Aufnahmebereich gleich. Gelegentlich kommen Supernieren (Supercardioid) oder Hypernieren (Hypercardioid) bei Sprachaufnahmen für Hörbücher zum Einsatz.⁶⁸

⁶⁶ Götze: 04.02.2022.

⁶⁷ Smyrek 2012: S.148.

⁶⁸ Götze 04.02.2022, siehe Anhang 1.

Schalleinfall	Spannung, Pegel bezogen auf die Spannung bzw. den Pegel des Schalls von vorne			
	Niere	Superniere	Hyperniere	Acht
von vorne (0°)	100 %, 0 dB			
seitlich (90°)	50 %, – 6 dB	38 %, – 9 dB	25 %, – 12 dB	0 %, – ∞ dB
von hinten (180°)	0 %, – ∞ dB	25 %, – 12 dB	50 %, – 6 dB	100 %, 0 dB
beste Eignung für die Aufnahmesituation	Schall von hinten soll ausgeblendet werden	Schall von der Seite und besonders von hinten soll ausgeblendet werden	Schall besonders von der Seite und von hinten soll ausgeblendet werden	Schall von der Seite soll ausge- blendet werden, Schall von vorn und hinten soll in gleicher Weise aufgenommen werden

Abbildung 17 Pegelverhältnisse aus verschiedenen Schalleinfallrichtungen von Richtmikrofonen⁶⁹

Um das Richtverhalten von Mikrofonen sinnvoll vergleichen zu können, sind Nierencharakteristiken standardisiert worden. Bei gleicher Richtcharakteristik sollen Mikrofone verschiedener Hersteller die gleiche richtungsabhängige Empfindlichkeit aufweisen. Die Standardisierung bietet einen hilfreichen Anhaltspunkt. Trotzdem gibt es geringfügige Unterschiede der Richtcharakteristiken.

Ein Polardiagramm (Polar Pattern) veranschaulicht die Richtcharakteristik. Für alle Einfallrichtungen zeigt das Richtdiagramm die Mikrofonempfindlichkeit in einer geschlossenen Kurve.⁷⁰ Die zentrale Schnittebene der räumlichen Richtcharakteristik liegt rotationssymmetrisch zu der senkrecht die Membran schneidenden Achse. Weil die Richtcharakteristik frequenzabhängig ist, finden sich im Datenblatt eines Mikrofons Richtdiagramme für verschiedene Frequenzen. Bei idealen Mikrofonen ist die Richtcharakteristik frequenzunabhängig, was sich in der Praxis schwer realisieren lässt. Die abgebildeten Kurven weichen von idealen Richtcharakteristiken ab. Alle Mikrofone neigen bei Frequenzen mit Wellenlängen in Kapselabmessung dazu Schall unregelmäßiger aufzunehmen. Das Richtdiagramm von Mikrofonen mit Nierencharakteristik weist Entartungen bei tiefen Frequenzen auf. Bei Kleinmembranmikrofonen sind die Richtcharakteristiken oft frequenzunabhängiger als bei Großmembranmikrofonen. Ein Beispiel dafür ist die K4 Kleinmembrankapsel von Schoeps.⁷¹ Kapseln mit fest eingestellter Richtcharakteristik entsprechen verglichen mit umschaltbaren Mikrofonkapseln eher den idealen Richtcharakteristiken. Je unabhängiger die Richtcharakteristik ist, desto neutraler klingt das Mikrofon in Bezug auf ein diffuses Schallfeld. Ergänzend zu den Polardiagrammen sind Angaben der richtungsabhängigen Frequenzgänge sinnvoll. Ein Vergleich zwischen richtungsbezogenen Kurven, die die Mikrofonempfindlichkeit über das hörbare Frequenzspektrum abbilden veranschaulicht, wie Diffusschall und Direktschall aufgenommen werden.⁷²

⁶⁹ Dickreiter 2014: S.152.

⁷⁰ Smyrek 2012: S.375.

⁷¹ Götze 04.02.2022, siehe Anhang 1.

⁷² Dickreiter 2014: S.140.

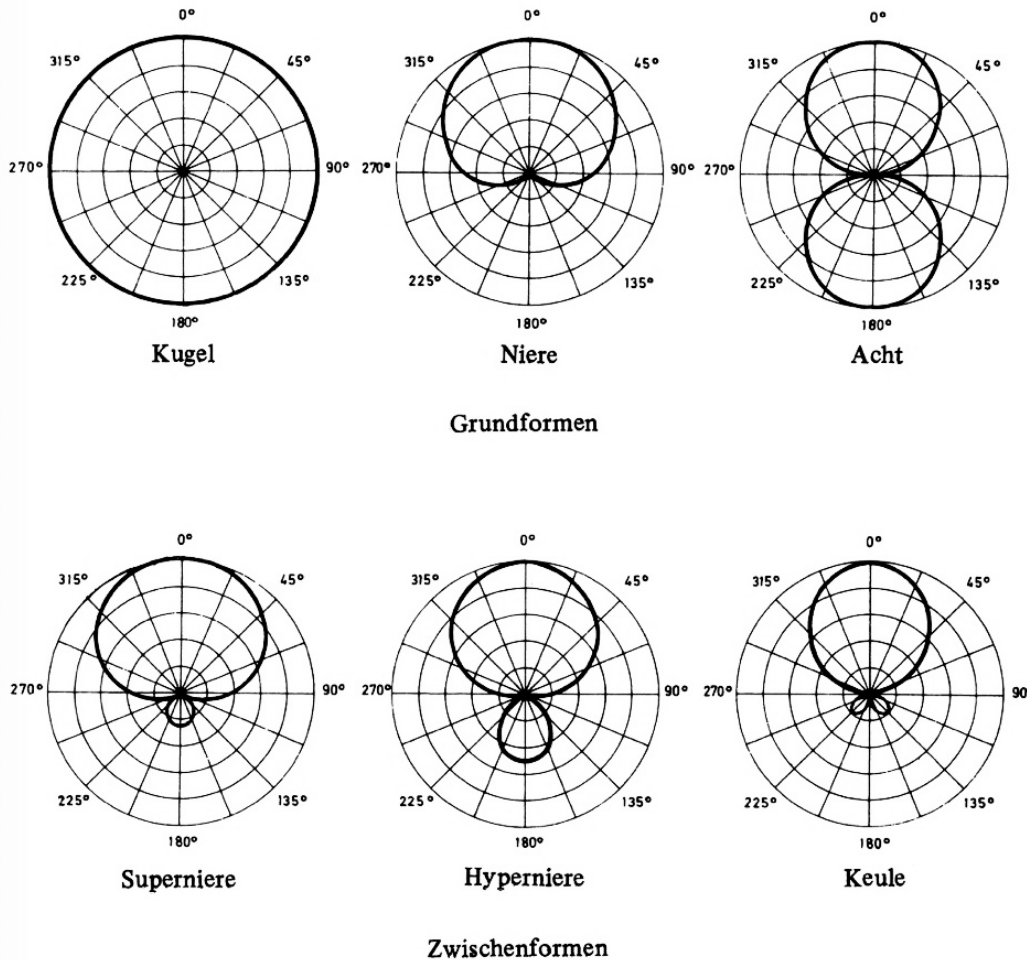


Abbildung 18 Polardiagramme von idealisierten Richtcharakteristiken⁷³

Im Polardiagramm wird das Richtungsmaß angegeben, das sich aus dem 20-fachen Zehnerlogarithmus des Richtungsfaktors berechnet und in dB angegeben wird. Der dimensionslose Richtungsfaktor hängt von der Schalleinfallrichtung und Frequenz ab. Er errechnet sich für eine bestimmte Frequenz und Richtung aus dem Verhältnis des Feld-Übertragungsfaktors für eine einfallende ebene Schallwelle aus einer bestimmten Richtung, zum Feld-Übertragungsfaktor aus der Bezugsrichtung. Bei letzterem handelt es sich um den Faktor der Bezugsachse oder Symmetrieachse (0°).⁷⁴ Der Feld-Übertragungsfaktor bei einer bestimmten Frequenz ist das Verhältnis zwischen der effektiven Ausgangsspannung und dem effektiven Schalldruck am Ort des Mikrofons, wenn dieses aus dem Schallfeld entfernt wurde. Falls keine explizite Frequenzangabe für den Feld-Übertragungsfaktor existiert, wird er für 1 kHz in mV/Pa angegeben, wobei 1 Pa einem Schalldruckpegel von 94 dB SPL entspricht. Je höher der Feld-Übertragungsfaktor eines Mikrofons ist, desto empfindlicher ist das Mikrofon.⁷⁵

⁷³ Dickreiter 2014: S. 151.

⁷⁴ Dickreiter 2014: S.148-149.

⁷⁵ Dickreiter 2014: S.140.

4.1.1 Bündelungsgrad und Bündelungsmaß

Mit den Kenngrößen Bündelungsgrad und Bündelungsmaß kann das Verhältnis von Direktschall zu Diffusschall verschiedener Mikrofone verglichen werden. Das Bündelungsmaß (Directivity Index) ist der zehnfache Logarithmus des Bündelungsgrades.⁷⁶ Der Bündelungsgrad eines Mikrofons gibt Auskunft darüber, um wie viel höher die Leistung des aufgenommenen Raumschalls bei demselben Mikrofon mit Kugelcharakteristik wäre. Der Feld-Übertragungsfaktor des Mikrofons bleibt bei der Angabe des dimensionslosen Bündelungsgrades gleich. Ein Mikrofon mit Kugelcharakteristik hat automatisch den Bündelungsgrad eins. Bei einem Mikrofon mit Nierencharakteristik erreicht nur 1/3 des Raumschalldrucks das Mikrofon aus der 0°-Richtung. Hätte das Mikrofon bei gleichem Feld-Übertragungsfaktor eine Kugelcharakteristik, würde es dreimal mehr Raumschallleistung aufnehmen. Somit ist der Bündelungsgrad für das Mikrofon drei. Wenn man statt der Schallleistung den Schallpegel betrachtet, ist der Besprechungsabstand beim Nierenmikrofon um $\sqrt{3}$ -mal größer als bei einem Kugelmikrofon (Vergrößerungsfaktor). Folglich hat der Raumschall erst bei 1,73 m den gleichen Einfluss auf das Nierenmikrofon wie bei dem Kugelmikrofon bei einem Meter. Das Bündelungsmaß berechnet sich aus dem 10-fachen Zehnerlogarithmus des Bündelungsgrades und wird in dB angegeben. Je größer der Bündelungsgrad oder das Bündelungsmaß ist, desto gerichteter ist ein Mikrofon. Infolgedessen können Mikrofone mit kleinem Bündelungsmaß in geringerem Abstand zum Sprecher platziert werden, während ein Mikrofon mit höherem Bündelungsmaß weiter weg stehen kann. Bei dieser Platzierung bleibt die Räumlichkeit ähnlich. Da bei der Hörbuchproduktion hauptsächlich Nierencharakteristiken eingesetzt werden, lohnt sich der Vergleich zwischen Niere, Super- und Hypernieren.

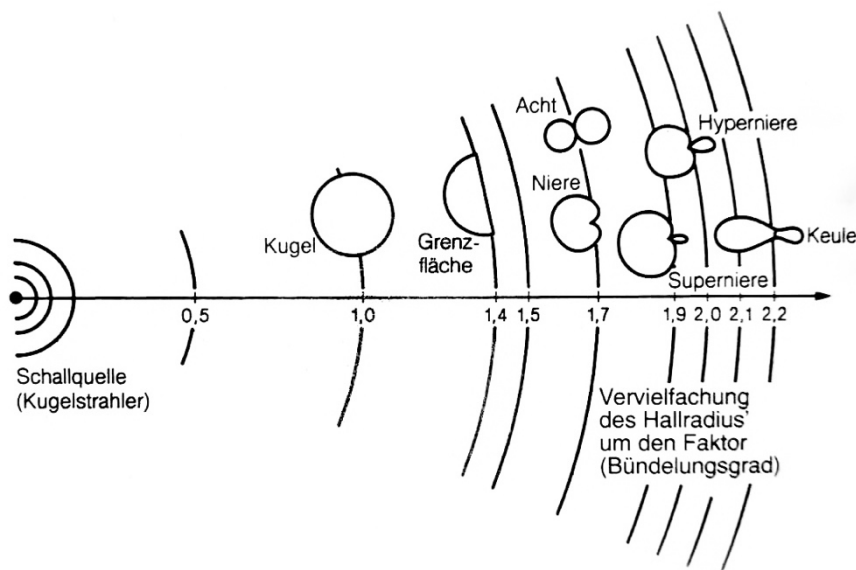


Abbildung 19 Mikrofonrichtcharakteristiken und ihre Vergrößerungsfaktoren im Vergleich⁷⁷

⁷⁶ Smyrek 2012: S.376-377.

⁷⁷ Dickreiter 2003: S.28.

4.1.2 Nahbesprechungseffekt

Der Nahbesprechungseffekt (Proximity Effect) wird bei Hörbuchaufnahmen als gestalterisches Mittel eingesetzt.⁷⁸ Bei Druckgradientenempfängern, also bei Richtmikrofonen mit Nieren- oder Achtrichtcharakteristik und deren Zwischenformen, tritt er im Nahfeld auf. Ab einem Abstand von ein bis eineinhalb Metern zwischen Mikrofon und Schallquelle setzt der Naheffekt ein.⁷⁹ Im direkten Nahfeld einer Schallquelle weisen diese Mikrofone aus physikalischen Gründen einen anderen Frequenzgang als für weiter entfernte Standorte auf. Der Schalldruck verhält sich im Nahfeld wie im Fernfeld. Allerdings steigt die Schallschnelle im Nahfeld viel stärker als der Schalldruck, wenn die Funktionsweise von Mikrofonkapseln auf der Schallschnelle oder dem Prinzip der Schalldruckdifferenz basiert.

Der Druckgradient steigt im Vergleich zum Schalldruck überproportional an, weil bei geringeren Abständen zur Schallquelle die Krümmung der Wellenfronten im Vergleich zur Wellenlänge zunimmt. Neben dem durch die Kapselbauweise bedingten Umweg verursachen Kugelwellen im Nahfeld eine zusätzliche Phasenverschiebung. Die Druckdifferenz zwischen Membravorder- und rückseite ergibt sich aus dem frequenzabhängigen Druckgradienten und der entfernungsabhängigen Schallabnahme. Aus dem Zusammenspiel von Nah- und Fernfeldkomponenten resultiert ein erhöhter Membranantrieb. Dieser ist umso größer, je kleiner der Abstand zwischen Schallquelle und Mikrofon ist und je tiefer die Frequenz ist. Eine Pegelanhebung setzt ungefähr da ein, wo der Mikrofonabstand kleiner als die Wellenlänge des auftreffenden Schalls wird.⁸⁰ Für hohe Frequenzen ist das Nahfeld so klein, dass der Nahbesprechungseffekt in der Praxis im hohen Frequenzbereich nicht auftritt. Aufgrund der zur Wellenlänge proportionalen Ausdehnung des Nahfeldes tritt der Pegelanstieg für tiefe Frequenzen bei Entfernungen bis zu einer Wellenlänge wirksam auf.⁸¹

Je mehr die Kapselbauweise eines Mikrofons einem Druckempfänger ähnelt, desto weniger macht sich der Nahbesprechungseffekt bemerkbar. Die Bassanhebung ist bei Hypernieren- und Achtmikrofonen besonders stark ausgeprägt. Bei Mikrofonen mit Nierencharakteristik tritt der Effekt allerdings auch deutlich hörbar auf. Bei weniger als 30 bis 40 cm werden tiefe Frequenzen eindeutig hörbar angehoben.⁸² 66 Hz werden bei einem kritischen Abstand von 35 cm um 3 dB angehoben und auf 131 Hz kann der Nahbesprechungseffekt bis zu einem Abstand von 17 cm wirken.

⁷⁸ Götz: 05.02.2022, siehe Anhang 1.

⁷⁹ Henle 2001: S.157-158.

⁸⁰ Dickreiter 2014: S.143-145.

⁸¹ Görne 2002: S.45-47.

⁸² Smyrek 2012: S.147-149.

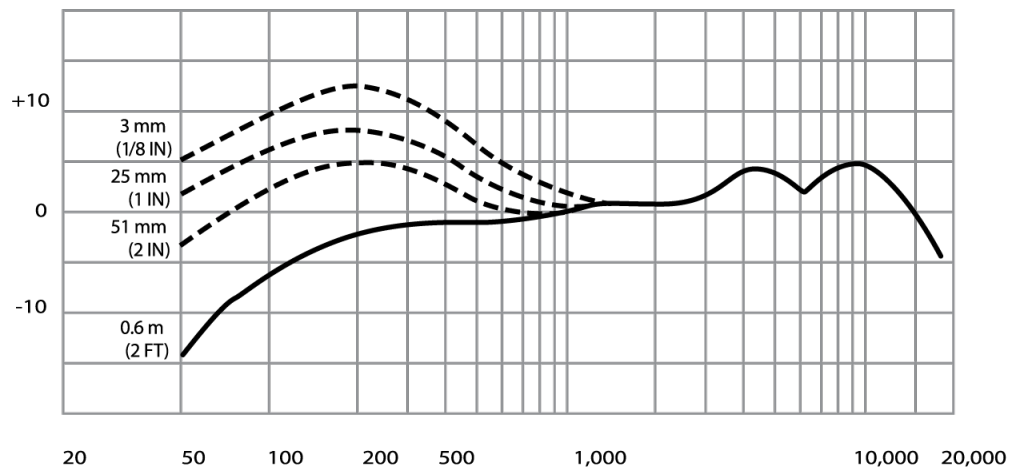


Abbildung 20 Entfernungabhängiger Frequenzgang einer Superniere beim Nahbesprechungseffekt. x-Achse: Frequenz in Hz, y-Achse: Schalldruckpegel in dB ⁸³

Trotzdem ist bei sehr kleinen Aufnahmeabständen Vorsicht geboten. Da der Nahbesprechungseffekt nicht nur frequenz- sondern auch entfernungsabhängig ist, sind sich ändernde Aufnahmeabstände durch den Naheffekt hörbar. Der Nahbesprechungseffekt kann durch eine festeingebaute oder schaltbare Bassdämpfung (Roll Off) kompensiert werden. Dicht an der Schallquelle könnten viele Mikrofone ansonsten nicht eingesetzt werden. Soll ein Gradientenempfänger neutral und möglichst signalgetreu aufnehmen, empfehlen sich ein- oder mehrstufige Schalter für verschiedene Einsprechabstände oder ein fester Mindestabstand, der größer als 0,5 bis ein Meter zwischen Mikrofon und Sprecher ist. Um den Nahbesprechungseffekt zu vermeiden kann der Gradientenempfänger durch einen Druckgradientenempfänger ersetzt werden. Mikrofone mit nativer Kugelcharakteristik weisen keine Bassanhebung im Nahfeld auf. Bei Kugelcharakteristiken von Mikrofonen mit umschaltbarer Richtcharakteristik kann es trotzdem zu einem Nahbesprechungseffekt kommen. Die mechanisch schaltbaren Kapseln sind Gradientenempfänger, auch wenn das umschaltbare Mikrofon auf „Kugel“ gestellt ist.

Nicht immer verbessert der Naheffekt die Sprachaufnahme. Der Effekt trägt zu einem wärmeren Stimmklang bei, sollte aber nur an inhaltlich angemessenen Stellen eingesetzt werden. Besonders nah aufgenommene Sprache klingt mit der Bassanhebung intimer. Bei männlichen Sprechern ist der Nahbesprechungseffekt bei sehr geringen Mikrofonabständen auffällig. In der Werbung oder in Kinotrailern wird der Effekt in der Regel in hohem Maße eingesetzt. Bei Sprachaufnahmen für Hörbücher ist der Naheffekt erwünscht, da er eine positiv wahrgenommene Wirkung auf das Stimmvolumen hat. Bei Sprachaufnahmen für Hörbücher kann der Effekt der Stimme mehr Wärme verleihen. Laut dem Hörbuchproduzent Markus Götze eignet sich ein durchschnittlicher Abstand von 20 bis 30 cm zum Mikrofon. Je nach inhaltlicher Angemessenheit variieren Sprecher diesen Durchschnittsabstand und setzen den Nahbesprechungseffekt als künstlerisches Mittel ein. ⁸⁴

⁸³ Shure GmbH: BETA58A. Gesangsmikrofon, in: Shure Publications, 2022 [online] <https://pubs.shure.com/guide/BETA58A/de-DE> [08.02.22]

⁸⁴ Götze: 05.02.2022, siehe Anhang 1.

4.1.3 Umschaltbare Richtcharakteristiken

Mikrofone mit umschaltbarer Richtcharakteristik ermöglichen einen schnellen klanglichen Vergleich in Bezug auf die Richtcharakteristik. Je nach Modell können verschiedene Charakteristiken und -mischformen eingestellt werden. Die Umschaltung erfolgt über einen Schalter am Mikrofon oder ferngesteuert.

Eine umschaltbare Richtcharakteristik bei Gradientenmikrofonen wird mit einer Doppelmembrankapsel umgesetzt, sodass sich auf beiden Seiten der Gegenelektroden eine Membran befindet. Im Prinzip handelt es sich um zwei Kapseln mit Nierencharakteristik, die zusammengefasst werden.⁸⁵ Über die elektronische Schaltung kann die Spannungsversorgung der Kapseln gesteuert werden. Wird nur eine Kapsel mit Spannung versorgt, erhält man eine Nierencharakteristik in Richtung der unter Spannung stehenden Kapsel. Wenn an beiden Kapseln die gleiche Spannung mit gleicher Polarität anliegt, arbeiten die Kapseln in alle Richtungen und es entsteht eine Kugelcharakteristik. Durch das Umpolen einer Kapsel ergeben sich zwei entgegengepolte Nierencharakteristiken, die sich zu einer Acht subtrahieren. Bei seitlichem Schalleinfall sind die Spannungen gegenphasig und heben sich deshalb komplett auf. Werden die Spannungsanteile bei der Achterrichtcharakteristik im Zweikapselsystem verändert, ergeben sich Zwischenformen wie eine Super- oder Hyperniere.⁸⁶

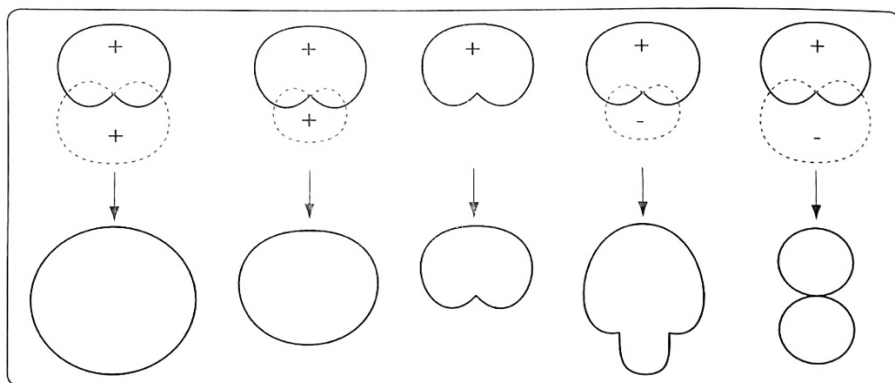


Abbildung 21 Richtcharakteristiken durch die Addition zweier Nieren⁸⁷

Im Hörbuchbereich ist der Vergleich zwischen verschiedenen Nierencharakteristiken attraktiv, weil hauptsächlich mit Nieren aufgenommen wird. In spezifischen Frequenzbereichen kann eine Richtcharakteristik im Vergleich zu einer anderen zu einem harmonischeren oder raumloseren Klang beitragen. An dieser Stelle sei erwähnt, dass man Richtcharakteristiken auch anhand verschiedener Mikrofone vergleichen kann, anstatt durch das Umschalten an einem Mikrofon. Allerdings tragen bei einem Vergleich mehrerer Mikrofone auch andere Kriterien abseits der Richtcharakteristik zum Klangbild bei.⁸⁸

⁸⁵ Boré 2002: S.37.

⁸⁶ Pawera 2003: S.47-49.

⁸⁷ Henle 2001: S.160.

⁸⁸ Götze: 05.02.2022, siehe Anhang 1.

4.1.3.1 Austrian Audio OC818

Das OC818 vom wiener Hersteller Austrian Audio ist ein Großmembran-Kondensatormikrofon mit umschaltbarer Richtcharakteristik. Klanglich orientiert es sich an den seit 1970ern von AKG Acoustics produzierten Mikrofonvarianten der Baureihe C414. Das Mikrofon verfügt über einen Low-Cut und Dämpfungsschalter. Ferner können mit einem Schalter die Richtcharakteristiken Acht, Superniere, Niere und Kugel eingestellt werden. Der Frequenzgang des Mikrofons in der Kugeleinstellung ist weitgehend linear und verhält sich nahezu wie ein Druckempfänger. Folglich tritt der Nahbesprechungseffekt nicht nennenswert auf. Die Kapsel des OC818 ist die Keramik-Doppelmembrankapsel CKR12, die auch als „Nylon Capsule“ bezeichnet wird. Bei dem neuen Kapseldesign kommen nur keramische Materialien für eine bessere Langzeitstabilität zum Einsatz. Akustisch orientiert sich die Kapsel an vor 1970 produzierten, linear abgestimmten, „Brass“-Kapseln. Das Klangbild des Mikrofons wirkt unabhängig von der Richtcharakteristik offen und natürlich. Obertöne der menschlichen Stimme werden auf angenehme Weise hervorgehoben, Zischlaute ziseliert und artefaktfrei übertragen. Dadurch klingen Höhen weicher.⁸⁹

In Bezug auf Richtcharakteristiken stellt das OC818 eine innovative Funktion zur Verfügung. Austrian Audio bietet ein kostenloses Plug-In, das eine nachträgliche Einstellung der Richtcharakteristik über bis zu fünf Frequenzbändern in der DAW erlaubt. Mit dem Plug-In PolarDesigner können neben Kugel, Niere und Acht sämtliche Zwischenformen eingestellt werden. Ein rückwärtiger Steckverbinder kann mit dem Bluetooth Dongle OCR8 belegt werden, über den sich die Richtcharakteristik auch per Smartphone über die App Polar Pilot steuern lässt. Vor allem das Plug-In zur Umschaltung der Richtcharakteristik ist sehr nützlich in der Postproduktion der Sprachaufnahmen.

Das OC818 eignet sich wegen der stufenlosen Änderung der Richtcharakteristik über mehrere Frequenzbereiche für Hörbuchaufnahmen im Tonstudio. Die Richtcharakteristik kann spontan analog oder in der Postproduktion digital umgeschaltet werden. Dadurch ist es dem Toningenieur möglich den Klang und die Räumlichkeit der Aufnahme nachträglich zu beeinflussen. In den einzelnen Frequenzbändern kann eine spezifische Richtcharakteristik einen angenehmeren Klang hervorbringen als eine andere. Insbesondere übermäßig betonte Laute, die in die entsprechenden Frequenzbereiche eingeordnet werden, können auf diese Weise korrigiert werden.⁹⁰

⁸⁹ Hau, Andreas: Austrian Audio OC818 Großmembran-Kondensatormikrofon im Test, in: Sound & Recording, 11.09.2020 [online] <https://www.soundandrecording.de/equipment/austrian-audio-oc818-grossmembran-kondensatormikrofon-im-test/> [16.11.21]

⁹⁰ Götze 04.02.2022, siehe Anhang 1.

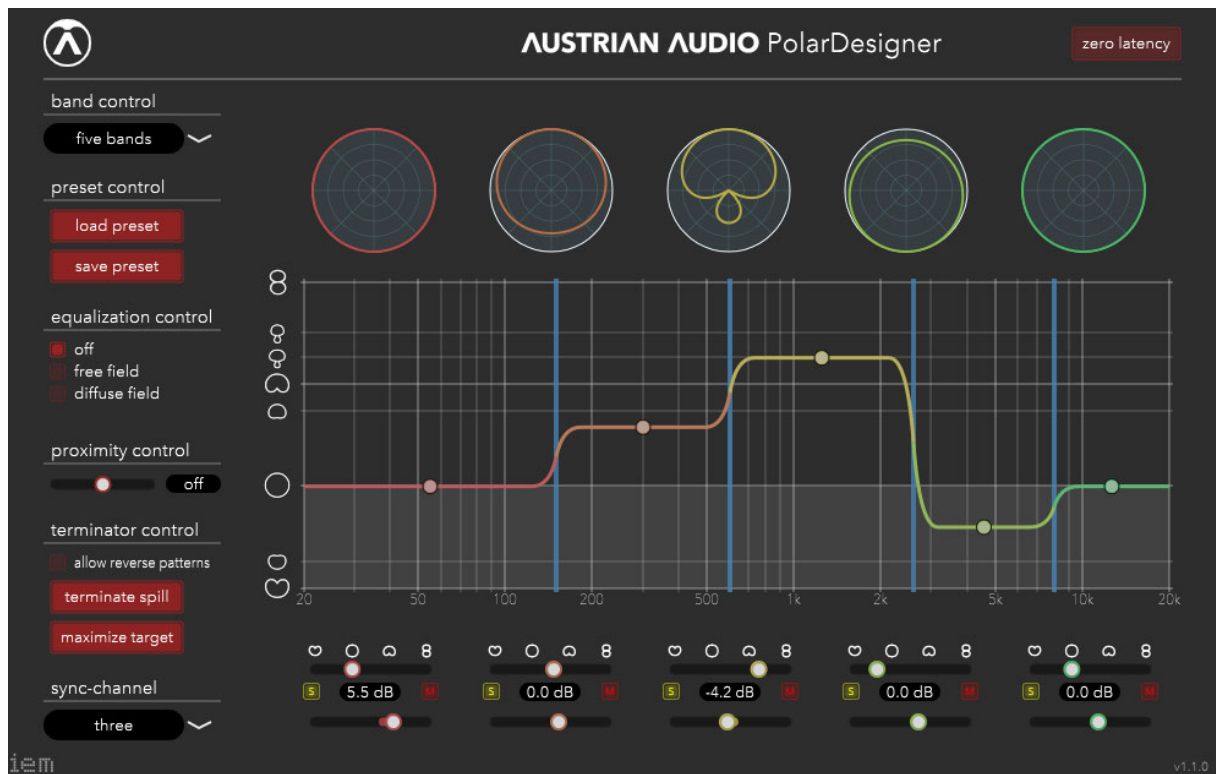


Abbildung 22 Plug-in zur Einstellung der Richtcharakteristik des OC818 in der DAW ⁹¹

4.2 Elektromechanische Wandlung

Die Elektromechanische Wandlung von Sprachsignalen erfolgt bei Sprachaufnahmen für Hörbuchproduktionen typischerweise elektrostatisch (dielektrisch). Kondensatormikrofone, die mit dem elektrostatischen Wandlungsprinzip funktionieren, eignen sich besonders für Sprachaufnahmen. Das liegt unter anderem an der guten Übertragungsqualität von Kondensatormikrofonen. Sprache wird verglichen mit dynamischen Mikrofonen neutraler und signalgetreuer abgebildet.⁹² In Abhängigkeit von Ausstattung und Konstruktion können Kondensatormikrofone über verschiedene klangliche Eigenschaften verfügen. Beispielsweise übertragen Kondensatormikrofone mit einer großen Membran einen wärmeren Klang als Kleinmembran-Kondensatormikrofone. Die stimmunterstützenden Konstruktionsweisen kann man sich bei Sprachaufnahmen zunutze machen.⁹³

⁹¹ AUSTRIAN AUDIO GmbH: Polar Designer und ihre kreative Freiheit, in: Austrian Audio, 2022 [online] <https://de.austrian.audio/polar-designer-und-ihre-kreative-freiheit/> [30.01.2022]

⁹² Görne 2002: S.186.

⁹³ Götze 04.02.2022, siehe Anhang 1.

Bereits 1928 entwickelte man elektrostatische Wandler für Kondensatormikrofone. In den 60er-Jahren erreichten Kondensatormikrofone mit dem CMV3 des Berliner Ingenieurs Georg Neumann eine Blütezeit, die bis heute anhält. Die eisenlos symmetrische Ausgangsschaltung stellt eine bedeutende Neuentwicklung der 80er-Jahre dar. Die durch die neue Technik gebauten transformatorlosen Mikrofone erlangen im Zuge der technischen Entwicklung immer mehr Beliebtheit. Die früher großen Mikrofone mit elektrostatischen Wandlern werden zunehmend kleiner, da sich die Halbleitertechnik und Mikroelektronik weiterentwickeln. Heute entsprechen Kondensatormikrofone aus den 50er- und 60er-Jahren immer noch dem zeitlosen Stand der Technik und sind wegen ihrer technischen Originalität begehrt.⁹⁴ Ältere Mikrofonmodelle können demnach genauso gut wie neue Mikrofone für Sprachaufnahmen verwendet werden.

Die Bezeichnung der Kondensatormikrofone lehnt sich an deren Funktionsweise an. Bei der elektromechanischen Wandlung werden die mechanischen Bewegungen der, von den Schallwellen in Bewegung gesetzten, Membran durch einen Plattenkondensator in elektrische Signale gewandelt. Der aus zwei Platten bestehende Kondensator ist über einen Schalter, einen Widerstand und eine Spannungsversorgung in einem elektrischen Schaltkreis eingebaut. Wenn sich der Schalter schließt, wird Energie aus dem Spannungsversorger entnommen. Dieser lädt den Kondensator in der sogenannten Ladezeit auf. Wenn der Kondensator aufgeladen ist, hört der Strom auf zu fließen und der Kondensator wirkt bei Gleichstrom als Isolator. Eine starre Metallplatte (Elektrode) und eine dünne elektrisch leitfähige Membran (Gegenelektrode) bilden beim Kondensatormikrofon die Platten des Kondensators.⁹⁵ Die Metallplatte ist als massiver, durchlöcherter Metallblock vorgesehen, der mit der Membran die Druckgradientenkapsel bildet. Der Kondensator ist so gebaut, dass ein Luftstrom möglich ist. Die Kapazität des Kondensators ändert sich mit der Membranbewegung, der periodischen Änderung des Elektrodenabstands. Bei größerer Kapazität nimmt der Kondensator mehr Ladung auf, während er bei kleinerer Kapazität eine kleinere Ladungsmenge absorbiert. Dabei entnimmt der Kondensator Energie aus der Batterie oder lädt sie auf. Folglich entsteht ein Stromfluss durch den Stromkreis. Die daraus hervorgehende Spannungsänderung entlang des Widerstandes entspricht dem elektrischen Signal der mechanischen Membranbewegung. So erzeugen Schallwellen ein Signal entlang des Widerstandes.⁹⁶

⁹⁴ Görne 2002: S.51-53.

⁹⁵ Webers, Johannes: Handbuch der Tonstudiotechnik für Film Funk und Fernsehen, 9. Auflage, Poing: Franzis, 2007.

⁹⁶ Henle 2001: S.143-147.

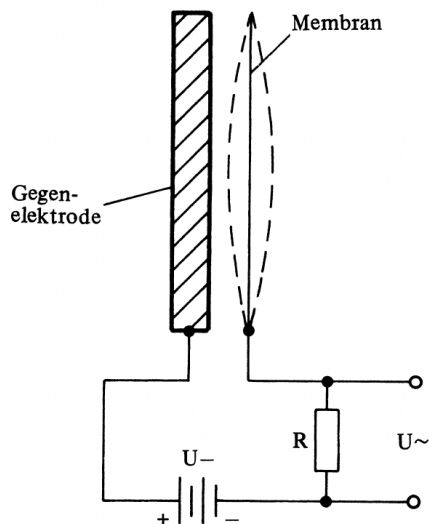


Abbildung 23 Schaltkreis einer elektrostatischen Wandlung ⁹⁷

Bei größerer Versorgungsspannung und Kondensatorkapazität ist der Abstand der Kondensatorplatten klein. Daraus folgt ein großer Strom durch den Stromkreis und eine hohe Mikrofonempfindlichkeit. In der Regel treten bei der elektromechanischen Wandlung in Kondensatormikrofonen Spannungen zwischen 50 und 100 V auf, wobei die Kapazität 20 bis 30 pF beträgt. Die Kapazität C_0 einer Kondensatormikrofonkapsel berechnet sich aus der Membranfläche S und dem Abstand zwischen Membran und der Metallplatte als Gegenelektrode d_0 .

Formel 4 Kapazität C_0

$$C_0 = \frac{\epsilon S}{d_0}$$

Für ein Großmembran-Kondensatormikrofon ergibt sich ein für eine Membranfläche mit 4,9 cm² und einem Elektrodenabstand von 0,04 mm eine Ruhekapazität von ungefähr 100 pF. Dabei ist ϵ die Dielektrizitätszahl der Luft. Wenn sich die Membran um einen Betrag d ändert, ändert sich die Kapazität C_0 um den Betrag wie folgt.

Formel 5 Kapazität C

$$C = \frac{C_0}{1 - \left(\frac{d}{d_0}\right)}$$

Die Kapazitätsänderung der Kapsel ist etwa proportional zur Membranauslenkung. Die Membran bewegt sich um 10 nm bei einem Pa Schalldruck. Wegen diesem Zusammenhang werden Kondensatormikrofone auch als Auslenkungs- oder Elongationswandler bezeichnet. Diese Funktionsweise ermöglicht es hochempfindliche Kondensatormikrofone mit sehr kleiner Membran zu bauen.⁹⁸

⁹⁷ Dickreiter 2014: S.138.

⁹⁸ Görne 2002: S.53.

Bei Kondensatormikrofonen wird nach der Kapsel ein impedanzwandelnder Vorverstärker eingebaut, weil die Kondensatorkapsel nur eine geringe Kapazität aufweist. Die Kapselspannung kann nicht störungsfrei als Ausgangssignal eines Mikrofons über einige Meter Kabel übertragen werden. Der Innenwiderstand der Kapsel ist zu hoch und ihre Kapazität zu klein.⁹⁹ Der Impedanzwandler (Spannungsfolger) passt den Innenwiderstand des Mikrofons durch einen großen Eingangs- und kleinen Ausgangswiderstand an die nachfolgende Verstärkerstufe an. Dabei wird die Spannung nicht verstärkt. Heutzutage ist die Impedanzwandlung mit einer Feldeffekttransistor-Schaltung (FET) gängig.¹⁰⁰ Bei den ersten Mikrofonen mit einer Niederfrequenzschaltung (NF) wurden Impedanzwandler in Röhrentechnik eingebaut. Angesichts der aus der Konstruktion hervorgehenden Klangeigenschaften sind Röhrenmikrofone heute noch beliebt.

Eine Kondensatorkapsel kann intern mit einer Niederfrequenzschaltung oder Hochfrequenzschaltung (HF) arbeiten. Die meisten Kondensatormikrofone werden in der NF-Technik gebaut. HF-Mikrofone sind in der Oberklasse der Studiomikrofone zu finden. Dass die interne Signalverarbeitung bei HF-Mikrofonen im hörbaren Frequenzbereich stattfindet, wird in der Nachrichtentechnik als Unterscheidung zwischen NF und HF angegeben. Es gibt dabei keinen nennenswerten Preis- oder Klangunterschied. Beide Varianten sollen das gleiche niederfrequente Signal am Ausgang des Mikrofons liefern.

Mit dem Frequenzbereich der Hochfrequenzschaltung ist der hochfrequente Bereich der Radioträgerwellen gemeint. Über einen sehr großen Ladewiderstand (hochohmige Kapsel) wird die Kondensatorkapsel in Reihe geschaltet. Durch eine externe Spannungsversorgung wird eine Kapselvorspannung erreicht, wodurch die Kondensatorkapsel auf eine definierte Gleichspannung (48 V, 60 V oder 200 V) aufgeladen wird. Bei Sprachaufnahmen im Tonstudio wird üblicherweise eine Phantomspeisung von 48 V an einem Vorverstärker eingeschaltet. Diese ist meist mit dem Kürzel P48 gekennzeichnet. Die Kapsel soll bei akustischer Anregung oberhalb einer bestimmten Grenzfrequenz keine wesentlichen Ausgleichsströme zwischen Membran und Gegenelektrode aufweisen. Entsprechend wird der Ladewiderstand definiert. Aus der Membranauslenkung resultiert die Kapselkapazität und ein proportionaler Spannungsabfall am Kondensator. Diese kann am Widerstand abgegriffen werden. Daraus folgt, dass die Nutzspannung bei großer Kapselspannung hoch ist.¹⁰¹

⁹⁹ Wirsum, Siegfried: Mischpulte und Mischpultmodule. Tonquellen, Mischpulte und Bausteine – Anleitungen für den Selbstbau nach Maß, 7. Auflage, München: Franzis 1988.

¹⁰⁰ Webers 2007: S.249.

¹⁰¹ Görne 2002: 56-58.

Das niederfrequente Nutzsignal wird bei der Hochfrequenzschaltung durch eine Modulationsschaltung in einen hochfrequenten Bereich geschoben. HF-Kondensatormikrofone sind mit einem HF-Oszillator parallelgeschaltet. Die Frequenz des Oszillators ändert sich mit der durch die Membranbewegung hervorgerufenen Kapazitätsänderung. Dadurch entsteht ein frequenzmoduliertes Signal (FM). Es ist durch eine alternative Schaltung auch eine Phasenmodulation (PM) oder Amplitudenmodulation (AM) möglich. Letztere wird in den modernsten Mikrofonen mit symmetrisch aufgebauten Kapseln verbaut. Der Innenwiderstand eines HF-Mikrofons ist klein. Für die Rückwandlung des FM-Signals in ein NF-Signal folgt auf den Oszillator ein Demodulator. Bevor das NF-Signal das Kondensatormikrofon verlässt, wird es von einem Verstärker bearbeitet. Bei HF-Mikrofonen wird keine Versorgungsspannung für die Kapsel gebraucht. Lediglich die HF-Schaltung benötigt eine Vorspannung. Bei HF-Mikrofonen wird der resultierende Frequenzgang zusätzlich elektrisch entzerrt.¹⁰²

4.3 Frequenzgang und Übertragungsbereich

Neben der Richtcharakteristik ist der Frequenzgang und der Übertragungsbereich eine wichtige Kenngröße von Mikrofonen. Bei Sprachaufnahmen für Hörbücher im Tonstudio entscheidet man sich für Mikrofone mit einem für den Sprecher passenden Frequenzgang und Kondensatormikrofone mit einem breiten Übertragungsbereich.¹⁰³

Der Frequenzgang gibt die frequenzabhängige Empfindlichkeit eines Mikrofons an. Bei idealen Mikrofonen sollte der Frequenzgang nicht richtungsabhängig sein. Der für eine Aufnahme nutzbare Frequenzbereich ist der Übertragungsbereich.¹⁰⁴ Begrenzt wird der Übertragungsbereich an denjenigen Frequenzen, bei denen der Ausgangspegel des Mikrofons im Vergleich zu dem Pegel bei 1 kHz um 3 dB kleiner ist. Hersteller missachten diese Studionorm allerdings oft, sodass beispielsweise der -8 dB - Übertragungsbereich stattdessen angegeben sein kann. Über den Übertragungsbereich geben Mikrofonhersteller Übertragungseigenschaften und deren Toleranzen an. Aufgrund unterschiedlicher Toleranzangaben können Mikrofone oft nicht direkt miteinander verglichen werden. Typische Übertragungsbereiche von Studiomikrofonen sind 20 Hz bis 20 kHz oder 40 Hz bis 16 kHz. Frequenzen unterhalb des Übertragungsbereich werden mit einem Hochpass weggefiltert, weil sehr tiefe Frequenzen Störungen im hörbaren Frequenzspektrum erzeugen können. Die Dämpfung sollte bei 40 Hz größer als 20 dB sein und bei 15 Hz ist vorzugsweise ein Pegelabfall von mindestens 12 dB zu verzeichnen.¹⁰⁵

¹⁰² Boré 2002: S.43-44.

¹⁰³ Götze: 04.02.2022, siehe Anhang 1.

¹⁰⁴ Smyrek 2012: S.153-154.

¹⁰⁵ Dickreiter 2014: S.141.

Der Frequenzgang wird auch Frequenzkurve genannt. Graphisch dargestellt zeigt die Frequenzkurve das frequenzabhängige Übertragungsmaß bei senkrecht auf der Membran auftreffendem Schall. Bei Studiomikrofonen darf der echte Frequenzgang um 2 dB von der angegebenen Kurve bei 1 kHz abweichen. Der Frequenzgang geht aus dem Zusammenwirken von Kapselkonstruktion und der elektromechanischen Wandlung hervor.

Mechanische Frequenzabhängigkeiten der Kapselkonstruktion führen zu bestimmten Resonanzfrequenzen. Eine schwingende Membran mit einer bewegten Masse m , einer Aufhängung mit der Federsteife s und die restliche Kapsel ist ein schwingungsfähiges System (Masseschwinger). Seine Resonanzfrequenz f_m steht unter Reibungsvernachlässigung in folgender mathematischer Beziehung.

Formel 6 Resonanzfrequenz f_m

$$f_m = \frac{1}{2\pi} \sqrt{\frac{s}{m}}$$

Demnach verkleinert sich die Resonanzfrequenz, wenn die Membranaufhängung weicher beziehungsweise die Membranmasse größer wird. Der Proportionalitätsfaktor wird hierbei vorerst nicht beachtet. Aus dem mathematischen Zusammenhang folgt, dass Kapseln mit schwerer Membran und vergleichsweise schwacher Membraneinspannung eine tiefe Resonanzfrequenz aufweisen (tief abgestimmte Kapsel). Analog dazu verfügt eine hoch abgestimmte Kapsel über eine leichte, straff gespannte Membran. Die Resonanzfrequenz wird bei einer mittig abgestimmten Membran über den hörbaren Frequenzbereich gemittelt.¹⁰⁶

Damit die Resonanz in der Praxis unbemerkt bleibt und die Membran nicht nachschwingt, gibt es eine kapselseitige Resonanzdämpfung. Trotzdem bleibt die Membrangeschwindigkeit und Membranauslenkung frequenzabhängig. Beide Größen unterscheiden sich in Bezug auf den Frequenzgang maßgeblich. Bei einem idealen Druckgradientenempfänger liegt immer eine Membranbewegung vor. Allerdings ist die Membrangeschwindigkeit im nahezu stillen Raum sehr gering. Unter einem erhöhten statischen Luftdruck wäre ein Membranausschlag messbar, aber die Membranschnelle gleich Null. Die Membrangeschwindigkeit nimmt mit steigender Frequenz zu, während die Auslenkungsamplitude der Membran zu höheren Frequenzen hin konstant bleibt. Membranschnelle und Membranauslenkung erreichen unter der Resonanzfrequenz ihre Maxima. Oberhalb der Resonanzfrequenz fallen beide Größen proportional zur Frequenz ab. Der Auslenkungsfrequenzgang zeigt eine Tiefpasscharakteristik, während der Schnellefrequenzgang über eine Bandpasscharakteristik verfügt. Alle mechanischen Frequenzabhängigkeiten durch die Mikrofonskapsel wirken als Filter erster Ordnung. Mit anderen Worten können die Frequenzgänge der Membranauslenkung und Membrangeschwindigkeit mit einfachen elektrischen Schaltungen simuliert werden, die ein elektrisches Signal in Abhängigkeit der Frequenz manipulieren.¹⁰⁷

¹⁰⁶ Görne 2002: S.62-68.

¹⁰⁷ Friesecke, Andreas: Die Audio-Enzyklopädie. Ein Nachschlagewerk für Tontechniker, 2. Auflage, Berlin: De Gruyter Saur 2014.

Elektrische Frequenzabhängigkeiten ergeben sich aus dem Ladewiderstand R_L und der Kapazität der Kapsel C_0 . Von beiden Faktoren ist die Grenzfrequenz der Kapsel abhängig. Der Proportionalitätsfaktor ist in diesem Zusammenhang nicht bedeutend. Ein NF- und HF-Kondensatormikrofon kann vereinfacht als RC-Glied dargestellt werden.¹⁰⁸ Die Schaltung erzeugt einen Hochpassfrequenzgang erster Ordnung. Daraus ergibt sich, dass die elektrische Grenzfrequenz f_{RC} wie folgt berechnet wird.

Formel 7 Grenzfrequenz f_{RC}

$$f_{RC} = \frac{1}{2\pi} \cdot \frac{1}{R_L \cdot C_0}$$

Bei HF-Kondensatormikrofonen liegt eine modulierte Hochfrequenz im MHz-Bereich an. Dadurch ist die Hochpassgrenzfrequenz weit unterhalb des Arbeitsbereiches. Zu tiefen Frequenzen werden HF-Mikrofone mit elektrischen Filtern begrenzt, um Störsignale im Infraschallbereich zu vermeiden.

Durch die Kombination von einem elektromechanischen Wandler mit einer Kapsel überlagern sich die Frequenzgänge beider Komponenten. Die Kopplung wirkt in Bezug auf die Frequenzgänge wie eine Reihenschaltung von zwei Filtern. Damit geeignete Frequenzgänge gefunden werden können, muss das Wandlerprinzip miteinbezogen werden. Die Funktionsweise elektrostatischer Wandler basiert auf der Membranauslenkung. Deshalb überlagern sich bei Kondensatormikrofonen mit Gradientenempfängern der Auslenkungsfrequenzgang mit der Hochpasscharakteristik. Der Auslenkungsfrequenzgang muss bei Druckempfängern mit dem RC-Hochpass zusätzlich gekoppelt werden.

Auffällig sind die Ähnlichkeiten zwischen den Frequenzgängen von Kondensator-Druckempfängern und dynamischen Gradientenempfängern. Kondensator-Druckempfänger verfügen über einen linearen und breiten Übertragungsbereich, der von der Kapselresonanz und der elektrischen oder akustischen Grenzfrequenz bestimmt wird. Kondensator-Druckempfänger haben eine hoch abgestimmte Kapsel. Sie sind im Vergleich zu ihrem dynamischen Pendant (dynamischer Gradientenempfänger) akustisch empfindlicher und neigen weniger zu nichtlinearen Verzerrungen. Zu tiefen Frequenzen ist der Frequenzgang einzig durch die Dimensionierung der RC-Gliedes begrenzt. Wenn der Ladewiderstand R_L ausreichend tief festgelegt wird, kann ein Kondensatormikrofon sehr tiefe Frequenzen übertragen. Ein dynamischer Gradientenempfänger bietet im tieffrequenten Bereich weniger Spielraum, da die Kapsel nicht beliebig tief gestimmt werden kann. Darüber hinaus ist die obere Übertragungsgrenze bei Kondensator-Druckempfängern mit oft 20 kHz höher als bei dynamischen Druckempfängern. Letztere übertragen meistens bis 16 kHz. Durch einen Höhenresonator vor der Membran des dynamischen Druckempfängers wird der Frequenzgang nach oben hin erweitert.¹⁰⁹

¹⁰⁸ Friesecke 2014: S. 419-422.

¹⁰⁹ Görne 2002:S. 62-68.

Kondensator-Gradientenempfänger besitzen einen ebeneren Frequenzgang als dynamische Druckempfänger. Beide Kombinationen aus Kapsel und elektromechanischem Wandler verfügen über eine mittig abgestimmte Kapsel. Das hat zur Folge, dass der Frequenzgang dazu neigt mittenbetont zu sein. In der Praxis wird die mittige Resonanz mechanisch so gedämpft, dass der Frequenzgang deutlich linearer wird. Wenn akustische Schwinger oder elektrische an die Konstruktion angekoppelt werden, erweitert und glättet sich der Übertragungsbereich zusätzlich. Durch diese Optimierungen wird auch mit einer mittig abgestimmten Kapsel ein hochwertiger Frequenzgang erzeugt. Dennoch können Frequenzen unterhalb von 40 Hz schlechter als bei Kondensator-Druckempfängern übertragen werden, was durch den Nahbesprechungseffekt kompensiert werden kann.¹¹⁰

Signalgetreue Sprachaufnahmen im Tonstudio lassen sich durch einen unabhängig von Frequenz und Richtcharakteristik möglichst ebenen Frequenzgang realisieren. Die K4 Kleinmembrankapsel des Herstellers Schoeps verfügt beispielsweise über einen im Vergleich zu anderen Kapseln ausgeprägt linearen Frequenzgang.¹¹¹ Die unterschiedlichen Frequenzgänge von Mikrofonen können allerdings auch der klanglichen Gestaltung dienen. Der Frequenzgang eines Mikrofons verzerrt den Frequenzgang des originalen Sprachsignals.¹¹²

Aus diesem Grund bietet es sich an mithilfe eines geeigneten Frequenzganges Defizite des Sprechers zu korrigieren. Bei ungeeigneter Mikrofonwahl wird die Stimme unvorteilhaft verfärbt. Zum Beispiel eignet sich das Mikrofon U87 des Herstellers Neumann für männliche Sprecher oder tiefe Stimmen, weil dessen Frequenzgang höhere Frequenzen anhebt. Infolgedessen wird tiefen Stimmen mehr Brillanz verliehen. Aus demselben Grund eignet es sich jedoch nicht für höhere Frauenstimmen. Diese klingen durch das U87 spitz. Für den täglichen Studiobetrieb im Hörbuchbereich bietet sich ein Frequenzgang an, der universell funktioniert. Das OC818 des Herstellers Austrian Audio weist einen solchen Frequenzgang auf. Laut dem Hörbuchproduzent Markus Götze ist es zwar kein makellostes Mikrofon, aber es liefert verlässlich hochwertige Aufnahmen bei vielen Sprechern.¹¹³

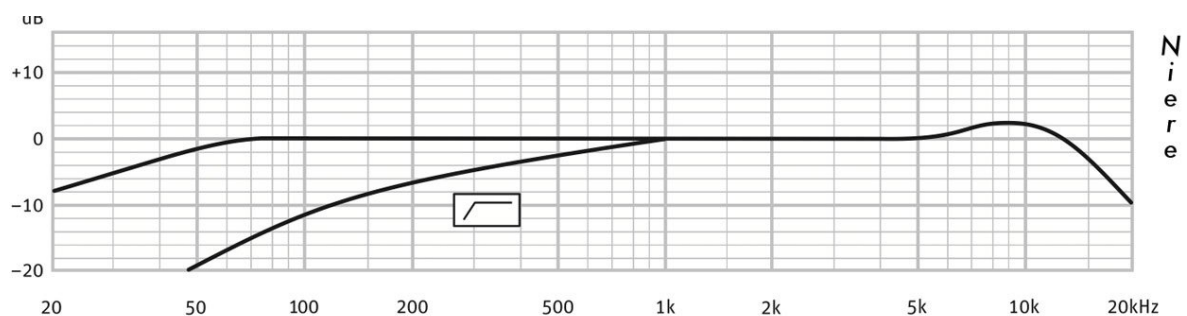


Abbildung 24 Höhenanhebung des U87 von Neumann mit und ohne Low-Cut ¹¹⁴

¹¹⁰ Görne 2002: S. 62-68.

¹¹¹ Götze: 04.02.2022, siehe Anhang 1.

¹¹² Boré 2002: S.65.

¹¹³ Götze: 04.02.2022, siehe Anhang 1.

¹¹⁴ Bauer, Armin: Test. Neumann U87 Ai, Grossmembran Mikrofon, in: AMAZONA.de music can change the world, 2022 [online] <https://www.amazona.de/test-neumann-u-87-ai-grossmembran-mikrofon/3/> [07.02.22]

4.4 Übertragungsfaktor und Übertragungsmaß

Die Größen Übertragungsfaktor und Übertragungsmaß beschreiben den Zusammenhang zwischen dem Schalldruck an der Mikrofonmembran und der Ausgangsspannung am Mikrofon. Weitere charakterisierende Maße für Mikrofone, wie Störspannungsangaben, beziehen sich auf die Größen. Der Übertragungsfaktor B gibt an, wie viel Spannung ein Mikrofon bei einem bestimmten Schalldruck abgibt. Die Wandlerempfindlichkeit (Sensitivity) wird so mit einem 1 kHz-Sinuston bei einem Schalldruck von einem Pascal (94 dB SPL) gemessen. Die Größe berechnet sich aus dem Quotienten der effektiven Ausgangsspannung und dem effektiven Schalldruck am Mikrofon, das aus dem Schallfeld entfernt wurde. Je größer der Übertragungsfaktor ist, desto empfindlicher ist das Mikrofon.¹¹⁵

Der auch als Feld-, Betriebs- oder Leerlauf-Übertragungsfaktor (Open Circuit Voltage) angegebene Wert wird in mV/ Pa oder in mV/ μ bar angegeben. Bei der Messung für den Übertragungsfaktor im Leerlauf misst man die Spannung an den offenen Mikrofonklemmen. Weiter verbreitet ist jedoch die Betriebsmessung, bei der das Anschließen eines Mikrofons an ein Mischpult simuliert wird. Dazu wird ein hochohmiger Widerstand an die Mikrofonklemmen geschaltet. Im Leerlauf sind die Spannungswerte höher als bei der Betriebsmessung, weil im Betrieb die Mikrofonklemmenspannung am Widerstand abfällt. Dadurch wird ein Teil der Spannung zur Verlustspannung. Somit simuliert die Betriebsspannung die durch den Innenwiderstand des Mikrofons und den Anschluss an das Mischpult abfallende Spannung. Für Kondensatormikrofone ist ein typischer Betriebsübertragungsfaktor 8 - 25 mV/ Pa. Darüber hinaus geben Kondensatormikrofone einen um 10 bis 20 dB höheren Pegel als dynamische Mikrofone ab. Existiert keine Angabe zur Messmethode des Übertragungsfaktors, sollte sicherheitshalber von einem im Leerlauf gemessenen Wert ausgegangen werden.

Es ist möglich Übertragungsfaktoren umzurechnen, damit zwei Mikrofone mit unterschiedlich gemessenen Faktoren verglichen werden können. Dazu wird zunächst die Spannungsverteilung u_{Betrieb} im Stromkreis berechnet. Die Spannung verteilt sich auf die Widerstände im Stromkreis. Es werden der Innenwiderstand des Mikrofons R_i (Nennimpedanz) und der Widerstand des Mischpultausgangs R_A als Widerstände des Stromkreises definiert. Die Spannung am Abschlusswiderstand R_A ist die Nutzspannung.¹¹⁶

Formel 8 Betriebsspannung u_{Betrieb}

$$u_{\text{Betrieb}} = R_A \cdot \frac{U_{\text{Leer}}}{R_i + R_A} \text{ mit } R_A = 1k\Omega$$

Das Übertragungsmaß ist der 20-fache Zehnerlogarithmus des Verhältnisses von einem Übertragungsfaktor zu einem Bezugsübertragungsfaktor. Letzterer liegt zum Beispiel bei 1 V/ Pa und wird in dB angegeben.

¹¹⁵ Henle 2001: S.162.

¹¹⁶ Görne 2002: S.171.

4.5 Nennimpedanz und Nennabschlussimpedanz

Schaltet man mehrere tonstudioteknische Geräte zusammen, hat das Verhältnis von Ausgangswiderstand eines Geräts zum Eingangswiderstand des angeschlossenen Geräts Einfluss auf die Verstärkung. Deshalb sind Bedingungen für die Zusammenschaltung erforderlich, die als Anpassung bezeichnet werden. Bei der Wahl des Mikrofons für die Sprachaufnahme muss die Nennimpedanz und Nennabschlussimpedanz geprüft werden, damit keine Störungen infolge der Zusammenschaltung von Mikrofon und Vorverstärker auftreten.

Der Nennwiderstand (Rating Impedance, Impedanz) ist der maximal zulässige, aber ungünstige Ausgangswiderstand (Quellwiderstand, Ausgangsimpedanz, Innenwiderstand). Dieser Scheinwiderstand wird bei 1 kHz gemessen und ist frequenzabhängig. Damit möglichst wenig Störungen auf Grund der Anpassung auftreten, sollte der Ausgangswiderstand eines Mikrofons so gering wie möglich sein.¹¹⁷ Der Nennwiderstand gibt an, ob es sich um ein nieder-, mittel- oder hochohmiges Mikrofon handelt. Professionelle Studiotechniker verwenden ausschließlich niederohmige Mikrofone, die nicht mehr als 200 Ohm betragen. Kondensatormikrofone besitzen einen spannungsversorgten Impedanzwandler, mit dem eine Impedanz von höchstens 40 Ohm realisiert wird. Durchschnittliche mittelohmige Mikrofone besitzen eine Nennimpedanz von bis zu 1 kOhm, während die Nennimpedanz von hochohmigen Mikrofonen bei ungefähr 100 kOhm liegt. Kondensatormikrofone verfügen über einen gleichmäßigen Impedanzverlauf.

Der Nennabschlusswiderstand (Load Impedance, Abschlussimpedanz) bezeichnet die für einen Betrieb gerade noch zulässige Belastbarkeit der Ausgänge. Der Widerstand gibt Auskunft darüber, wie ein Ausgang maximal belastet werden darf und liegt bei Mikrofonen mit höherem Ausgangswiderstand entsprechend höher. Alle Herstellerangaben zu Frequenzgängen und Eigenrauschen werden bei Widerständen, die größer als der angegebenen Eingangswiderstand des Verstärkers sind, eingehalten. Bei niederohmigen Mikrofonen ist standardmäßig 1 kOhm angegeben, was von dem Großteil der Mischpulthersteller eingehalten wird. Brauchen Mikrofone weniger als 1 kOhm, haben sie ein gutes Übertragungsverhalten bei verschiedenen Anschlussimpedanzen.

In der Praxis sollte bei einem auffallend hohen Anschlusswiderstand eines Mischpultes Vorsicht geboten sein, wenn ein sehr niederohmiges Mikrofon angeschlossen wird. Ist die Nennabschlussimpedanz weniger als 10- bis 20-mal höher als die Impedanz des Mikrofons, treten frequenzabhängige Auswirkungen auf den Frequenzgang des Mikrofonpegels auf.¹¹⁸ Bei einem unangemessenen Ausgangswiderstand am Vorverstärker wird der Klang des Mikrofons beeinträchtigt. Daher ist es wichtig, dass Verstärker und Mikrofon aneinander elektrisch angepasst sind.¹¹⁹

¹¹⁷ Dickreiter 2014: S.499-500.

¹¹⁸ Görne 2002: S.170-171.

¹¹⁹ Götz: 05.02.2022, siehe Anhang 1.

Der Eingangswiderstand sollte nach Vorgaben der öffentlich-rechtlichen Rundfunkanstalten mindestens zehnmal größer sein. Bei häufigen Eingangswiderständen von 200 Ohm sind mindestens 600 Ohm am Ausgangswiderstand gefordert. Dieser Wert wird von vielen Geräten überschritten. Praktisch können daher bedenkenlos zwei Geräte aneinander geschaltet werden. Erst bei mehr als zwei Geräten muss die Zulässigkeit geprüft werden.¹²⁰ Das Kapitel analoge Echtzeitbeeinflussung von Sprachsignalen beleuchtet technische Voraussetzungen für das Zusammenschalten analoger Geräte im Tonstudio.¹²¹

4.6 Störpegel

Bei der Wahl eines geeigneten Mikrofons für Sprachaufnahmen zählt der Störpegel zu den qualitätsbestimmenden Parametern. Er gibt Auskunft über die Übertragungseigenschaften eines Mikrofons. Unter völliger Stille liegt an den Mikrofonklemmen eine Spannung an. Dieses Eigenrauschen des Mikrofons wird über die Fremdspannung oder Rauschspannung angegeben. Alle aufnahmetechnischen Geräte rauschen, allerdings ist Mikrofonrauschen besonders kritisch. Es beeinflusst das Mikrofonsignal, wobei das Rauschen durch die dem Mikrofon folgende Verstärkung ebenfalls angehoben wird.

Nach der temperaturabhängigen Teilchenbewegungen (Bown'sche Bewegungen) stehen Luftmoleküle bei totaler Stille nicht still. Ihre ständige Bewegung erzeugt eine vom Widerstandswert abhängige Rauschspannung (Widerstandsrauschen, Wärmerauschen, Luftauschen). Es entsteht zudem Rauschen, da sich auch die Ladungsträger in den Dioden und Transistoren von Mikrofonen bewegen. Abhängig vom Konstruktionsprinzip sorgen weitere Schaltungen wie Filter und HF-Schaltungen in Kondensatormikrofonen für weitere Spannungen.¹²² Im Übertragungssystem wird durch elektrische Störquellen die Fremdspannung hervorgerufen. Aus ihnen resultiert ein akustischer Störpegel. Die Fremdspannung (unweighted) wird je nach Norm über einen großen Übertragungsbereich zwischen 22,4 Hz und 22,4 kHz gemessen. Dabei nähert sich die Messung im mittleren Bereich einer Geraden an und fällt zu den Grenzen des Übertragungsbereichs ab.¹²³

Da die Teilchenbewegung der Luftmoleküle mit steigender Temperatur schneller wird, tritt bei kälteren Temperaturen weniger Wärmerauschen auf. Bei 0 K (-273 °C) bewegen sich weder Luftmoleküle noch Elektronen. Deshalb wäre am absoluten Nullpunkt keine Spannung zu erwarten.¹²⁴ Demnach scheint es sinnvoll Mikrofone zu kühlen, um eine geringste mögliche Spannung zu erhalten. Weil die durch eine Kühlung verringerte Spannung unverhältnismäßig gering wäre, verzichtet man allerdings darauf. Eine Halbierung der Temperatur würde nur 3 dB weniger Rauschen zur Folge haben.¹²⁵

¹²⁰ Dickreiter 2014: S.499-500.

¹²¹ Görne 2002: S.170-171.

¹²² Görne 2002: S.173.

¹²³ Henle 2001: S.419-422.

¹²⁴ Görne 2002: S.173-174.

¹²⁵ Friesecke 2014: S. 213.

Das Wärmerauschen tritt zusammen mit dem akustischen Signal auf. Es wirkt qualitativ nach der Störwirkung in der Wahrnehmung durch das Gehör. Weil das menschliche Ohr nicht alle Störfrequenzen gleich laut empfindet, ist es sinnvoll die subjektive Lautstärkewahrnehmung (Kurven gleicher Lautstärke) in die Angabe des Störpegels miteinzubeziehen.¹²⁶ Die Geräuschspannung (Rauschspannung, bewertete Eigenstörspannung, weighted) berücksichtigt diese wertende Eigenschaft des Gehörs. In der Regel ist die Geräuschspannung kleiner als die Fremdspannung.¹²⁷

Praktisch ist der absolute Wert der Rauschspannung für eine Mikrofonauswahl wenig aussagekräftig. Stattdessen ist entscheidend, wie laut ein vom Mikrofon übertragenes Schallsignal relativ zum Eigenrauschen übertragen wird. Dieses Verhältnis wird vom Ersatzgeräuschpegel angegeben. Er ist eine von der Rauschspannung abgeleitete Größe. Hersteller geben die Rauschspannung von Mikrofonen in der Regel über diese Größe an. Mit dem Ersatzgeräuschpegel oder dem äquivalenten Schalldruckpegel (Equivalent Noise Level) wird die Hörbarkeit des Rauschens von Mikrofonen angegeben. Neben der Rauschspannung hängt der Messwert von der Mikrofonempfindlichkeit ab. Je größer der Übertragungsfaktor eines Mikrofons ist, desto größer ist das Nutzsignal im Verhältnis zum Störsignal. In diesem Fall wäre ein Mikrofon rauscharm. Ein kleiner Übertragungsfaktor führt zu stärkerem Rauschen.

Für die Berechnung des Ersatzgeräuschpegels geht man von einer Schallquelle vor dem Mikrofon aus, die allein für das Rauschen verantwortlich ist. Die subjektive Stärke des Rauschsignals ist der Schalldruckpegel dieser imaginären Schallquelle. Aus diesem Grund heißt die Messgröße auch Ersatzpegel. Diesen berechnet man, indem die bewertete gemessene Störspannung in mV durch den Übertragungsfaktor B in mV/ Pa teilt und in den hypothetischen Schalldruck umrechnet.

Formel 9 Rauschspannung L_{Rausch}

$$L_{Rausch} = 20 \cdot \log \left(\frac{u_{Rausch}}{B} \cdot \frac{1}{2 \cdot 10^{-5}} \right) dB SPL$$

Durch vorgeschaltete Filter, die sich an die Kurven gleicher Lautstärke anlehnen, wird das subjektive Hörempfinden des Menschen simuliert. Verschiedene Normen halten sich mit unterschiedlicher Genauigkeit an die Kurven. Daraus resultiert, dass sich die Angaben für Messwerte der Störpegel unterscheiden. Ein gängiger Bewertungsfilter ist der A-bewertete Ersatzgeräuschpegel, der den Effektivwert der Rauschspannung (RMS-Messung) angibt. Wird der Ersatzpegel nach CCIR-Norm gemessen, wird eine Spitzenwertmessung (Peak-Messung) durchgeführt. Der A-bewertete Pegel ist ungefähr 10 dB kleiner als der Ersatzpegel nach CCIR-Norm und seine Angabe deshalb von Vorteil für Hersteller.¹²⁸

¹²⁶ Smyrek 2012: S.156.

¹²⁷ Friesecke 2014: S.213.

¹²⁸ Görne 2002: S.174-175.

Ohne Angabe der Messmethode sollte vorsichtshalber davon ausgegangen werden, dass der Ersatzgeräuschpegel A-bewertet ist. Typische Werte nach A-Bewertung für Kondensatormikrofone sind 12-24 dB(A). Kondensatormikrofone rauschen wegen deren höheren Übertragungsfaktoren in der Regel weniger als dynamische Mikrofone.¹²⁹

Der Abstand des Ersatzgeräuschpegels zu einem Schallsignal mit 94 dB SPL (1 Pa) Schallstärke ist der Geräuschspannungsabstand (SNR, Signal-to-Noise Ratio). Er berechnet sich aus der Differenz von Geräuschspannungsabstand und Ersatzgeräuschpegel. Neben dem Ersatzgeräuschpegel, liegt mit dem SNR eine Aussage darüber vor, wie weit der Rauschpegel unter einem durchschnittlichen Nutzsignal ist. Es sei darauf hingewiesen, dass die englische Bezeichnung Signal-to-Noise Ratio häufig ohne Messmethode angegeben wird. Deswegen geht aus vielen Angaben, die einen sogenannten SNR aufführen nicht hervor, ob es sich um den Fremd- oder Geräuschspannungsabstand handelt.¹³⁰ Vergleicht man Mikrofone bezüglich ihrer Störpegel, ist darauf zu achten, um welche Größen genau es sich in den Datenblättern handelt. Ansonsten ist der Vergleich nicht aussagekräftig. Kleiner als 62 dB sollte der Geräuschpegelabstand von Mikrofonen für Sprachaufnahmen nicht sein. Bei einem geringeren Wert muss der Rauschbeitrag der nachfolgenden Geräte der Übertragungskette im Tonstudio bedacht werden. Generell bietet es sich an, ein Kondensatormikrofon für Hörbuchaufnahmen im Tonstudio zu verwenden. Bei hochwertigen Kondensatormikrofonen liegt der Geräuschpegelabstand bei mindestens 70 dB. Der Rauschbeitrag muss für den nachfolgenden Verstärker nicht beachtet werden, weil sie eine fünf- bis zehnmal so große Nutzspannung abgeben.¹³¹ Beispielsweise liegt der SNR bei dem OC818 von Austrian Audio bei mehr als 84 dB.¹³²

Rauschen sollte in der Praxis vermieden werden. In der Hörbuchproduktion entscheidet man sich für ein Mikrofon, das einen hohen Rauschabstand während der Aufnahme hervorbringt. Ist überdurchschnittlich viel Vorverstärkung notwendig, um einen ausreichenden Signalpegel zu erzeugen, rauscht das Signal. In diesem Zusammenhang ist zu erwähnen, dass ein rauschfreies oder -armes Signal aus einer angemessenen Aussteuerung am Vorverstärker hervorgeht. Wird das Signal nicht ausreichend verstärkt muss es nach der Aufnahme digital angehoben werden, was wiederum zu Rauschen führt.¹³³

¹²⁹ Görne 2002: S.174-175.

¹³⁰ Görne 2002: S.420.

¹³¹ Boré 2002: S.65.

¹³² Thomann GmbH: Austrian Audio OC818 Studio Set, in: Thomann, 2022 [online]
https://www.thomann.de/de/austrian_audio_oc818_studio_set.htm?gclid=CjwKCAiAo4OQBhBBEiwA5KWu_7vtNep0rzTJe1InmO5uQ9Q9JDA82D7U9guAshaJluEiKQUWD0tLhoCVUEQAvD_BwE [07.02.22]

¹³³ Götze 04.02.2022, siehe Anhang 1.

4.7 Impulsverhalten

Damit eine Aufnahme verglichen mit dem originalen Schallereignis einen identischen Klang aufweist, müsste das Impulsverhalten eines Mikrofons so sein, dass ein akustischer Schalldruckimpuls wie das Ausgangssignal (Impulsantwort) eines Mikrofons verläuft. Je nach Mikrofon weicht das Impulsverhalten allerdings vom Schalldruckimpuls ab. Eine lange Impulsantwort zeichnet sich durch einen unpräzisen, weichen Klang aus. Im Vergleich zu diesem schlechten Impulsverhalten beobachtet man bei kurzen Impulsantworten wenig Nachschwingen. Das Impulsverhalten von Mikrofonmembranen wirkt sich demnach auf die Klangcharakteristik von Mikrofonen aus.

Da Sprachsignale zahlreiche impulsive Laute enthalten, eignen sich Mikrofone mit einem möglichst guten Impulsverhalten für signalgetreue Sprachaufnahmen. In der Hörbuchproduktion werden häufig Großmembranmikrofone eingesetzt, da deren längeres Impulsverhalten der Sprachaufnahme eine warme Färbung verleiht. Ob sich ein Klein- oder Großmembranmikrofon besser für Sprachaufnahmen eignet, ist Geschmackssache. Mit letzterem wird die Aufnahme jedoch so gefärbt, dass sie sich eher nach einer typischen Hörbuchaufnahme anhört.¹³⁴ Das Impulsverhalten kommt nach der Theorie der linearen Systeme dadurch zustande, dass Frequenzgang und Impulsverhalten voneinander abhängen. Mikrofone können als solche linearen Systeme betrachtet werden. Nach der Theorie schwingt ein System bei breitem Übertragungsbereich kürzer nach, während es bei einem schmalbandigen Übertragungsbereich länger nachschwingt. Daraus resultiert, dass Kondensatormikrofone mit breitem Übertragungsbereich auf Grund der konstruktiv bedingt kleineren Bandbreite sauberer klingen.¹³⁵

Aus der Kapselkonstruktion lässt sich ebenfalls das Impulsverhalten ableiten. Das Impulsverhalten von Kondensatormikrofonen ist generell besser für das impulsive Sprachsignal geeignet. Bei Kondensatormikrofonen ist die Kapselresonanz mittig oder hoch, weshalb die Membranspannung nicht so nachgiebig wie die tief abgestimmte Kapselmembran von dynamischen Mikrofonen ist. Darüber hinaus ist die Membran des Kondensatormikrofons verglichen mit Membranen von dynamischen Mikrofonen leichter und schwingt darum nicht bedeutend nach.¹³⁶ Außerdem ist die Kapselimpedanz bei Kondensatormikrofonen hoch, was ein weniger verfremdetes Signal bei der Schallwandlung zur Folge hat.¹³⁷ Mechanisch ist die mittig abgestimmte Kapsel bei Kondensator-Gradientenempfängern äußerst stark gedämpft. Bei hoch abgestimmten Kondensator-Druckempfängern liegt die Eigenfrequenz der Kapsel an der oberen Grenze des Übertragungsbereichs. Die Membran ist hier sehr straff und die Impulsantwort noch kürzer als die der Kondensator-Gradientenempfänger.

¹³⁴ Götze 04.02.2022, siehe Anhang 1.

¹³⁵ Görne 2002: S.180-182.

¹³⁶ Henle 2001: S.163-165.

¹³⁷ Friesecke 2014: S.405.

Messungen zum Impulsverhalten werden von Herstellern selten bereitgestellt. Von der Konstruktions- und Funktionsweise eines Mikrofons kann man sich eine mehr oder weniger lange Impulsantwort jedoch ableiten. Ähnlich konstruierte Mikrofone können durch Aufnahmen desselben Schallimpulses miteinander verglichen werden.

4.8 Maßnahmen gegen Störschall

Störgeräusche entstehen, weil die Mikrofonmembran auf jegliche Art bewegter Luft reagiert. Luftwirbel und Winddruck gegen die Membran führen zu sogenannten Windstörungen. Schwingungen, die über bauliche Elemente übertragen werden erzeugen Körperschall. Dieser überträgt sich auf die Mikrofonmembran und bewegt sie obendrein. Beide Formen von Störschall sollen bei Sprachaufnahmen für Hörbücher im Tonstudio vermindert werden.

Im Nahfeld entstehen Übersteuerungen durch impulsivartige Luftströmungen bei Popp-, Zisch- oder Explosivlauten.¹³⁸ Bei Popp- oder Ploppstörungen wird Schall in eine definierte Richtung abgestrahlt. Die Luft strömt gleichgerichtet aus dem Mund und geht mit zunehmender Entfernung ab 10 cm in Windströmungen über. Im gerichteten Luftstrom kann das Mikrofon aus dem Wind gedreht werden. Wenn Wind verwirbelt aus allen Richtungen auf das Mikrofon trifft, kann es nicht aus dem Wind gedreht werden. Wind- und Poppstörungen weisen hohe Pegel im unteren Frequenzbereich bis ungefähr 100 Hz auf. Da die Störungen auch im Infraschallbereich auftreten, führen Ploppstörungen zu starken Verzerrungen. Infolgedessen wird das Mikrofon „zugestopft“. Damit ist gemeint, dass Bauelemente des Mikrofons übersteuern und das Signal in Pegelspitzen unterbrochen wird.¹³⁹

Konstruktionsbedingt reagieren Gradientenempfänger besonders empfindlich auf Luftbewegungen. Die Wind- und Poppempfindlichkeit (Wind and Pop Sensitivity) kann als A- oder CCIR-bewerteter Schalldruckpegel angegeben werden. Sie liegt bei Druckgradientenempfängern in einer Größenordnung von 20 dB.¹⁴⁰

¹³⁸ Boré 2002: S.49.

¹³⁹ Dickreiter 2014: S.163.

¹⁴⁰ Görne 2002: S.180.

Anblas- und Windgeräusche müssen deshalb durch zusätzliche Maßnahmen abgeschwächt werden. Ein Wind- oder Poppschutz bewirkt eine hierfür geeignete Abschirmung.¹⁴¹ Bastler können eine solche Vorrichtung aus einem Holz- oder Metallring bauen und mit zwei Lagen Gaze oder Nylonstrumpf bespannen. Beide Lagen sollten etwa einen halben bis ganzen Zentimeter auseinander liegen. Im Handel werden auch Ausführungen aus Schaumstoff angeboten. Allerdings setzt man für Hörbuchaufnahmen die ringförmige Vorrichtung zum Großteil ein. Eine bewegliche Halterung am Ring (Schwanenhals) erleichtert die Positionierung der Schutzvorrichtung. Das Gewebe kann zur Reinigung abgenommen werden. Dadurch, dass Luftströmungen bei einem Wind- oder Ploppschutz umgelenkt auf das Mikrofon treffen, wird die Windgeschwindigkeit vermindert. Folglich wirken sich Poppstörungen gering auf das Nutzsignal ein. Die Schutzvorrichtung hat Einfluss auf das Richtverhalten und den Frequenzgang eines Mikrofons. Im Tonstudio werden Poppschirme gegen Poppstörungen eingesetzt, da sie die technischen Daten des Mikrofons nur geringfügig verfälschen. Sprache wird für Hörbücher statisch aufgenommen, weshalb sich der Einsatz eines Poppschirms mit einem Stativ eignet.¹⁴²

Eine weitere Art der Störung wird als Körperschall oder Trittschall bezeichnet. Schwingungen pflanzen sich über den Boden oder Griffbewegungen auf den Mikrofonständer fort.¹⁴³ Das Mikrofon wird relativ zur Membran bewegt, was sich auf die Membranbewegung zur Aufnahme auswirkt. Praktisch wirken Mikrofonspinnen oder andere federnde Mikrofonhalterungen Körperschall entgegen. Oft wird die für das Mikrofon passende Spinne vom Hersteller mitgeliefert. Die elastische Aufhängung entkoppelt das Mikrofon von Erschütterungen. Außerdem verursachen Reibegeräusche am Mikrofonkabel bei Gradientenempfängern Körperschall. Deswegen sollten Kabel stets mit einer Zugentlastung am Stativ geführt werden. Einige Spinnen verfügen für diesen Zweck über eine Kabelkammer. Bei von der Decke hängenden Mikrofonen ist die Hängevorrichtung so zu bauen, dass eine Zugentlastung gewährleistet ist. Stative mit körperschallentkoppelnden Ausstattungen (Trittschallstativ) wie Gummifüße verhindern, dass sich Trittschall vom Boden auf Mikrofonständer überträgt. Hochwertiges und langlebiges Mikrofonzubehör trägt zur Qualität der Sprachaufnahme bei. Der Hersteller König & Meyer setzt den Standard für Stative und Schutzvorrichtungen, die sich für Sprachaufnahmen im Studio eignen.¹⁴⁴ Insbesondere der Mikrofonarm 23860 des genannten Herstellers bietet eine geeignete Halterung für Hörbuchaufnahmen am Lesetisch.¹⁴⁵ Tieffrequenter Tritt- und Körperschall kann ebenfalls durch einen Low-Cut vermindert werden.¹⁴⁶

¹⁴¹ Götze 04.02.2022, siehe Anhang 1.

¹⁴² Görne 2002: S.147-150.

¹⁴³ Pawera 2003: S.169.

¹⁴⁴ Görne 2002: S.147-150.

¹⁴⁵ König & Meyer: 23860 Mikrofonarm, in: König & Meyer. Stands for Music, 2022 [online] <https://www.km.de/produkte/mikrofonstative/mikrofonarme/23860-mikrofonarm-schwarz?c=128> [11.02.2022]

¹⁴⁶ Götze 04.02.2022, siehe Anhang 1.

4.9 Mikrofon positionieren

Die Position und der Schalleinfallswinkel des Mikrofons bestimmen die Räumlichkeit und den Klang der Sprachaufnahme. Aus einer Sprachaufnahme für Hörbücher soll eine möglichst „trockene“ Aufnahme hervorgehen, durch die der Klang der Stimme positiv beeinflusst wurde. Die Sprecherstimme bei Hörbüchern soll möglichst dicht am Zuhörer erklingen. Es wird danach gestrebt nur Direktschall und möglichst wenig Diffusschall aufzunehmen. Das heißt, dass einzig die Sprecherstimme ohne raum- oder effektbedingten Hall aufgenommen werden soll. Die aufgenommene Sprache soll in der Kopfmittle zu verorten sein. Weil es sich außerdem um eine einzige aufzunehmende Schallquelle handelt, reicht es aus, ein Mikrofon für die Aufnahme zu positionieren. Es wird Mono aufgenommen. Für eine Auswahl aus mehreren Mikrofonen können zusätzlich weitere aufgestellt werden. Im Hörbuch ist allerdings später nur ein Mikrofonsignal von einem der aufgestellten Mikrofone zu hören.

Es sei erwähnt, dass alle notwendigen raumakustischen Anforderungen von dem für Sprachaufnahmen optimierten Aufnahmerraum im Tonstudio erfüllt werden. Die Schallschutzkabine wie die STUDIOBOX ist eine für diesen Zweck geeignete Sprecherkabine. Ohne zusätzliche Baumaßnahmen kann die Kabine im Tonstudio platziert werden. Dabei bleibt die Qualität der Schalldämmung zur äußeren Umgebung hin bei unterschiedlichen Studioräumen nahezu gleich. Die innere Akustik des „Raumes im Raum“ wird reguliert und bietet so eine optimierte Aufnahmeumgebung.¹⁴⁷

Durch Mikrofonaufstellung dicht an der Schallquelle wird die Räumlichkeit der Aufnahme reduziert.¹⁴⁸ Für ein direktes Sprachsignal sollte die Sprachaufnahme im Nahfeld stattfinden. Das Nahfeld endet bei einem Meter Abstand von der Schallquelle zum Mikrofon.¹⁴⁹ Bei Sprachaufnahmen für Hörbücher wird in der Regel ein Abstand zwischen 20 und 30 cm gewählt.¹⁵⁰ Je nach gewünschter Klangfärbung wird der Aufnahmeabstand definiert. Dieser beeinflusst den Stimmklang bei Gradientenempfängern maßgeblich. Durch den Nahbesprechungseffekt werden tiefe Frequenzen in Abhängigkeit von der Richtcharakteristik und dem Abstand zwischen Sprecher und Mikrofon verstärkt. Daraus resultiert ein wärmerer und vollerer Klang. Extrem nah aufgenommene Stimme suggeriert Intimität, klingt jedoch bei übertriebenem Naheffekt unnatürlich. Wenn bei dichter Mikrofonierung kein Nahbesprechungseffekt erwünscht ist, kann dieser durch eine Kugelcharakteristik oder Filterung der tiefen Frequenzen kompensiert werden. Im Kapitel Nahbesprechungseffekt wird die Tiefenanhebung durch dieses Phänomen erläutert.

¹⁴⁷ STUDIOBOX GmbH: STUDIOBOX Schallschutzkabine, in: STUDIOBOX Schalldämmende Akustik Studios, 2022 [online] <https://www.studiobox.de/> [11.01.22]

¹⁴⁸ Webers 2007: S.166.

¹⁴⁹ Görne 2002: S.193.

¹⁵⁰ Götze 04.02.2022, siehe Anhang 1.

Bei ungünstiger Mikrofonaufstellung können Interferenzen die Aufnahme verfärben, insbesondere wenn sich mindestens zwei gleichförmige Signale auf die Mikrofonmembran treffen. Bei Sprachaufnahmen für Hörbücher kommt es zu keiner Überlagerung durch den Einsatz mehrerer Mikrofone, weil man sich für ein Mikrofonsignal entscheidet. Bei der Sprachaufnahme mit einem Mikrofon, interferieren dennoch Signale. Im Kapitel Kammfiltereffekt wird die klangliche Verfärbung aufgrund der Signalüberlagerung erklärt.

Schallreflexionen an Wänden lassen sich durch eine Mikrofonposition vermeiden, die eine maximale Entfernung zu den Flächen garantiert. Das Mikrofon sollte sich also ungefähr in der Raummitte befinden. Bei einem Mikrofon mit Nierencharakteristik, kann das Mikrofon näher zur Wand auf der Mikrofonrückseite platziert werden. Während Sprachaufnahmen für Hörbücher sitzt der Sprecher in der Regel an einem Tisch. Der Tisch dient der Ablage des Lesemediums. Anstelle einer Ablage kann der zu lesende Text auf einem Bildschirm angezeigt werden. In beiden Fällen ruft die zusätzliche Fläche in unmittelbarer Nähe des Mikrofons Schallreflexionen hervor. Daher ist es ratsam, das Mikrofon in größtmöglicher Distanz zur Lesefläche zu positionieren. Darüber hinaus kann es in einem Winkel gehängt werden, der es erlaubt, dass die Richtcharakteristik des Mikrofons die Lesefläche weitestgehend ausblendet. Alternativ zur hängenden Mikrofonposition, ist eine dicht über dem Tisch stehende Platzierung denkbar. Zeigt dieses zur Schallquelle und verfügt über eine Richtcharakteristik, welche die rückseitige Schallaufnahme verhindert, werden Interferenzen vermieden.¹⁵¹

In der Praxis entscheidet man sich häufig für ein hängendes Mikrofon, das etwa 20 - 30° über der horizontalen Blickrichtung auf den Mund des Sprechers zeigt. So hindert das Mikrofon den Sprecher nicht am Lesen. Außerdem treffen Zischlaute und Plosive nicht auf direktem Weg auf die Mikrofonmembran, da sie vom Sprecher aus gesehen nach vorne unten abgestrahlt werden. Folglich ist ein positiver Nebeneffekt eines hängenden Mikrofons von oben ein weicherer Klang mit weniger aggressiven Zischlauten und abgeschwächten Plosiven. Bei Sprache mit starken Zischlauten kann das Mikrofon so geneigt werden, dass es etwas seitlich zum Mund steht. Damit wird ein noch weicherer Klang erzielt.¹⁵²

Fehler bei der Mikrofonaufstellung können zu einer qualitativ schlechteren Aufnahme führen. Zwar ist es möglich diese mit neuester Software zu vermindern, jedoch können sie nicht vollständig behoben werden. Insbesondere die audiorestaurierenden Plug-Ins aus der RX-Reihe des Herstellers iZotope führen zu erstaunlichen Ergebnissen. Die Räumlichkeit lässt sich so in der Postproduktion vermindern, allerdings geht dieser Arbeitsschritt meist mit hörbaren Artefakten einher.¹⁵³ Aus diesem Grund wird postproduktionsseitiges Bearbeiten der Räumlichkeit vermieden und stattdessen eine korrekte Mikrofonierung angestrebt. Generell gilt je besser die Position des Mikrofons ist, desto weniger Nachbearbeitung für aufnahmeseitige Korrekturen folgt.

¹⁵¹ Görne 2002: S.206-207.

¹⁵² Friesecke 2014: S.157.

¹⁵³ Thornton, Mike: iZotope RX 9 Tested - The Good, The Bad And The Ugly, in: PRO TOOLS EXPERT, 14.10.21 [online] <https://www.pro-tools-expert.com/production-expert-1/izotope-rx-9-tested-the-good-the-bad-and-the-ugly> [11.01.21]

4.9.1 Kammfiltereffekt

Bei Sprachaufnahmen für Hörbücher im Tonstudio entstehen störende Klangverfärbungen, wenn sowohl der Direktschall vom Sprecher als auch Schallreflexionen aus dem Raum aufgenommen werden. Dieses Phänomen wird als Kammfiltereffekt (comb filter) bezeichnet. Insbesondere das Lesemedium und der Lesetisch stellen im Tonstudio für Hörbuchaufnahmen reflektierende Flächen dar. Allerdings sind beide Komponenten für die Aufnahme notwendig. Entsteht durch die Klangverfärbungen im Sprachsignal eine Kammfilterkurve, sind die Klangverfärbungen auffällig.¹⁵⁴

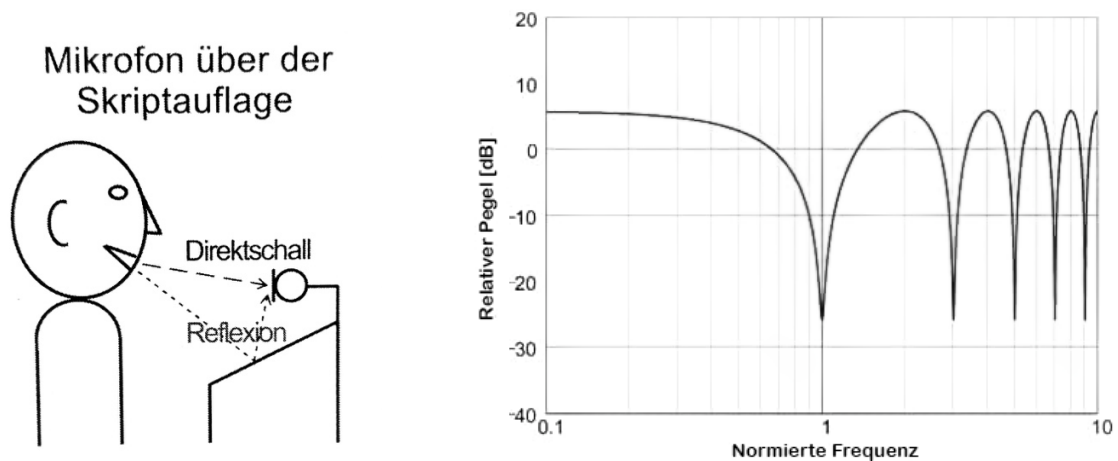


Abbildung 25 Kammfiltereffekt im Raum (links)¹⁵⁵ und resultierender Frequenzgang (rechts)¹⁵⁶

Der reflektierte Schall addiert sich in der Aufnahme zum Direktschall. Dabei entstehen Interferenzen. Es entsteht ein Zeitunterschied Δt zwischen dem Direktsignal und der Reflexion, da die Reflexion immer später auf die Mikrofonmembran trifft als das direkte Sprachsignal. Der Zeitunterschied ist abhängig von der Frequenz, weshalb verschieden große Phasenunterschiede zwischen dem Direktschall und der Reflexion entstehen. Aus der Frequenz f und dem Zeitunterschied Δt lässt sich der Phasenwinkel φ zwischen den Signalen in Grad berechnen.¹⁵⁷

Formel 10 Phasenwinkel φ zwischen direktem und reflektiertem Signal

$$\varphi = \Delta t \cdot f \cdot 360^\circ$$

¹⁵⁴ Dickreiter 2003: S.188.

¹⁵⁵ Smyrek 2012: S.179.

¹⁵⁶ Maier, Peter/ Lembke, Stephan: Raumakustik. Das Studio, der Eierkarton, die Absorber und Architektur, in: SOUND & RECORDING, 2022 [online] <https://www.soundandrecording.de/tutorials/raumakustik-das-studio-der-eierkarton-die-absorber-und-architektur/> [08.02.22]

¹⁵⁷ Friesecke 2014: S.29-30.

Ist die Phasendifferenz der Signale 180° oder ein geradzahliges Vielfaches davon, löschen sich Direktschall und reflektierter Schall bei einer Addition gegenseitig aus. Dabei kommt es nicht immer zu einer kompletten Auslöschung, weil das reflektierte Signal in der Regel leiser als das direkte Signal ist. Wenn beide Signale gleich laut sind und eine Phasendifferenz von 360° oder ein Vielfaches davon besteht, erhöht sich der Pegel bei einer der addierten Signale um 6 dB.¹⁵⁸

Die Auswirkungen auf den gesamten Frequenzbereich ähneln den Zählen eines Kamms. Es treten immer abwechselnd ein Einbruch und eine Überhöhung auf. Je nach Raumreflexion verschieben sich die Einbrüche und Überhöhungen entlang des Frequenzbereichs. Werden mehrere Frequenzen bei einer Addition von direktem und reflektiertem Schall ausgelöscht oder im Pegel erhöht, ist meist kein Kamm mehr erkennbar. Infolgedessen wird der Frequenzverlauf stark verzerrt. Schon eine Zeitverzögerung beider Signale zwischen einer und 15 ms führt zu Klangverfärbungen.¹⁵⁹ Abhängig von der Stärke des Kammfilters entstehen hörbare Verfärbungen in der Sprachaufnahme, die nicht rückgängig gemacht werden können. Deshalb sollte er durch entsprechende Maßnahmen vermieden werden. Bei der Mikrofonierung ist darauf zu achten, dass der Lesetisch möglichst ausgeblendet wird. Das Mikrofon ist so zu platzieren, dass es möglichst wenig Reflexionen aufnimmt. Deshalb eignet sich eine Mikrofonierung von oben, und leicht seitlich, sodass es auf den Mund des Sprechers gerichtet ist. Alternativ positioniert man das Mikrofon auf dem Lesetisch, wobei es zum Sprecher zeigt. Auch ein stärker gerichtetes Mikrofon als eine Niere, etwa eine Hyperniere, kann einen starken Kammfiltereffekt minimieren.¹⁶⁰



Abbildung 26 Kammfiltereffekt vermeiden: geeignete Mikrofonposition für eine Hörbuchaufnahme mit Maßnahmen gegen Störschall

¹⁵⁸ Smyrek 2012: S. 9.

¹⁵⁹ Dickreiter 2003: S. 14-16.

¹⁶⁰ Götze 04.02.2022, siehe Anhang 1.

5. Analoge Echtzeitbeeinflussung von Sprachsignalen

Zur Aufbereitung des Mikrofonsignals können unterschiedliche Geräte in den Signalweg eingebaut werden, die in Echtzeit wirken. Dieses Kapitel beschäftigt sich mit der analogen Audiotechnik, die bei Sprachaufnahmen für Hörbücher im Tonstudio eingesetzt wird. Ziel der analogen Echtzeitbearbeitung ist es die Nachbearbeitung der Sprachaufnahmen für Hörbücher zu minimieren. In diesem Zusammenhang ist zu erwähnen, dass die analoge Aufbereitung der Aufnahme nicht rückgängig gemacht werden kann. Alle Auswirkungen auf die Sprache werden mit aufgenommen. Aus diesem Grund muss bedacht vorgegangen werden. Bei angemessenem Umgang mit analogen Geräten kann schon während der Aufnahme ein verdichtetes Sprachsignal aufgenommen werden.

Bei den Sprachaufnahmen im Tonstudio wird die vom Schallwandler aufgenommene Sprache über ein XLR-Kabel, gegebenenfalls über ein Steckfeld, geleitet. Das Signal passiert anschließend analoge Geräte wie einen Mikrofonvorverstärker, einen Equalizer und einen Kompressor und eventuell einen Limiter. Danach wird das analoge Signal in ein digitales Signal gewandelt und zu diesem Zweck in ein Interface geleitet. Dieses ist mit einer digitalen Workstation (DAW) auf einem Computer verbunden. Alle Aufnahmen werden mithilfe der DAW gespeichert.

In der analogen Echtzeitbeeinflussung von Signalen wird die ursprüngliche Schallinformation in einem Format verarbeitet, sodass die aufgenommenen Schallereignisse durch Parameter verändert werden können. Bei der analogen Echtzeitbeeinflussung werden die Parameter so gewählt, dass die in eine elektrische Spannung gewandelte Schallinformation im Sinne der Hörbuchproduktion positiv beeinflusst wird.

Alle aneinander geschalteten analogen Geräte in der Übertragungskette müssen elektrisch aneinander angepasst sein (Anpassung).¹⁶¹ Bei der Anpassung müssten bestimmte Bedingungen für das Verhältnis zwischen dem Ausgangswiderstand und dem Eingangswiderstand nachfolgender Geräte erfüllt werden. Im Kapitel Nennimpedanz und Abschlussnennimpedanz wurden Ein- und Ausgangswiderstände in Bezug auf Mikrofone erklärt. Beim korrekten Zusammenschalten von mehr als zwei analogen Geräten sind folgende zusätzlichen Kenntnisse notwendig. Zwei zusammengeschaltete Geräte können als Spannungsteiler dargestellt werden.¹⁶² Das Spannungsverhältnis einer Leerlaufspannung U_0 zur Eingangsspannung U_E entspricht dem Verhältnis der Summe aus dem Eingang- und Ausgangswiderstand zum Eingangswiderstand.¹⁶³

¹⁶¹ Götze 04.02.2022, siehe Anhang 1.

¹⁶² Dickreiter 2014: S.498-500.

¹⁶³ Hufschmid, Markus: Grundlagen der Elektrotechnik. Einführung für Studierende der Ingenieur- und Naturwissenschaften, Wiesbaden: Springer Vieweg 2021.

Formel 11 Spannungsverhältnis einer Leerlaufspannung U_0 zur Eingangsspannung U_E

$$\frac{U_0}{U_E} = \frac{R_A + R_E}{R_E}$$

Sind Eingangs- und Ausgangswiderstand identisch, liegt eine Leistungsanpassung vor. Keine der übertragenen Leistungen ist maximal. Ein exakt linearer Frequenzgang ist bei denselben Werten für Eingangs- und Ausgangswiderstände nur mit großem Aufwand zu realisieren. Alle Geräte müssen für die gleiche Leistung dimensioniert sein. Zudem steigt der elektrische Widerstand (Scheinwiderstand) frequenzabhängig, wenn Kondensatoren und Spulen verbaut sind. Die Wirkung dieses Phänomens ist besonders bei hohem Stromfluss stark. Eine Kompensation des Frequenzgangs mittels Schaltungen ist so aufwändig, dass die Leistungsanpassung im Tonstudio für Aufnahmen keine große Rolle spielt. Sie beschränkt sich auf Endstufen und Laufsprecher.

In der Tonstudioteknik arbeitet man in der Regel mit der Stromanpassung (Überanpassung), um einen möglichst linearen Frequenzgang zu erreichen. Durch die Überanpassung können an einen Ausgang mehrere Eingänge angeschlossen werden, ohne dass das Ausgangssignal hinsichtlich Frequenzgang oder Pegel beeinflusst wird.¹⁶⁴ In der Theorie können bis zu zehn Geräte problemlos aufeinander folgen. Allerdings ist es unmöglich an einen Eingang eines Geräts zwei Ausgänge anzuschließen. Grund dafür ist der Ausgangswiderstand, der für das andere Gerät wie ein Eingangswiderstand wirkt. Eine Überanpassung liegt nicht mehr vor. Es folgen lineare und nicht-lineare Verzerrungen.¹⁶⁵

Bei der Stromanpassung ist der Ausgangswiderstand (Quellwiderstand) kleiner als der Eingangswiderstand. Der Strom wird in der Praxis wie bei einem Kurzschlussbetrieb übernommen. Eine große Anzahl an parallelgeschalteten Quellen bei dieser Anpassungsart ist möglich, ohne dass sich Tonsignale wechselseitig beeinflussen. Deswegen wird die Stromanpassung hauptsächlich in der Tonstudioteknik eingesetzt. Tonsignale lassen sich mit der Spannungsanpassung optimal zusammenführen. Dieses Verfahren bezeichnet man als Nullohm-Knotenpunktschaltung. Alle Ausgangswiderstände der zusammengeschalteten Geräte und deren Entkopplungswiderstände stellen in diesem Fall den Ausgangswiderstand R_A dar.

¹⁶⁴ Smyrek 2012: S.195-196.

¹⁶⁵ Henle 2001: S.67-69.

In der Praxis wird die Spannung in ihrer Originalhöhe weitergegeben. Es genügt, wenn der Ausgangswiderstand R_A kleiner als ein Zehntel des Eingangswiderstandes R_E ist. Dabei sind keine genauen Werte, sondern Mindest- oder Höchstwerte zu definieren. Demnach können mehrere Geräte zusammen geschaltet werden, wenn der Gesamtwiderstand nicht geringer als der Mindestwert wird. Der Ausgangswiderstand liegt bei tontechnischen Geräten bei maximal 40 Ohm. Die Ausgänge sind belastbarer, wenn ihr Quellwiderstand entsprechend höher ist. Die gerade noch zulässige maximale, aber ungünstige Belastbarkeit am Geräteausgang wird als Nennabschlussimpedanz bezeichnet. Bei 40 Ohm Ausgangswiderstand ist die Belastbarkeit etwa 300 Ohm. Ein Eingangswiderstand von mindestens 600 Ohm wird von den öffentlich-rechtlichen Rundfunkanstalten angegeben, der praktisch von dem meisten Geräten übertroffen wird.¹⁶⁶ Somit können viele Geräte bedenkenlos aneinander geschaltet werden.¹⁶⁷ In der Praxis wird ein Verhältnis von mindestens 25:1 angestrebt.¹⁶⁸

Die Anpassung der Übertragungskette im Tonstudio kann mithilfe eines Sinustons als Testsignal überprüft werden. Alle Geräte der Übertragungskette sind aneinander angepasst, wenn 0 dB am Vorverstärkerausgang nach der Kette auch 0 dB sind. Für Hörbuchaufnahmen ist das insofern bedeutend, dass wenn ein Mikrofon mit der Aussteuerung -6 dB angeschlossen wird, es in der DAW auch -6 dB sein sollten. Ansonsten sind die abzulesenden Schalldruckpegel unzuverlässige Werte.¹⁶⁹

Die analogen Komponenten der Übertragungskette sind jeweils auf einen bestimmten Arbeitspegel ausgelegt. In einem Pegeldiagramm wird der Pegel über die Übertragungskette dargestellt. Viele professionelle Mikrofonvorverstärker wandeln das Mikrofonsignal so um, dass es sich bei einem Arbeitspegel von 0 dBr, also 6 dB einpegelt. In der ganzen Übertragungskette wird dieser Arbeitspegel beibehalten. Alle externen Geräte sind dann in einer hochwertigen Übertragungskette auf den Arbeitspegel von 6 dB ausgelegt. Es gibt auch Übertragungsketten, bei denen der interne Arbeitspegel -10 dBr beträgt. Dieser Pegel hat den Vorteil, dass eine höhere Übersteuerungsreserve erreicht wird und sich der Schaltungsaufwand verringert. Aus letzterem ergeben sich weniger Kosten. Bei dieser semiprofessionellen Variante müssen alle an ein Mischpult angeschlossenen Geräte über einen Arbeitspegel von 0 dBr verfügen. Deshalb schaltet man einen Aufholverstärker ans Ende der analogen Signalbearbeitungskette. Auf den Aufholverstärker kann verzichtet werden, wenn generell mit einem geringeren Arbeitspegel gearbeitet wird. Bei hochwertigen Sprachaufnahmen im Tonstudio wird ein Arbeitspegel von 6 dB angestrebt.¹⁷⁰

¹⁶⁶ Dickreiter 2014: S.498-500.

¹⁶⁷ Dickreiter 2014: S.498-500.

¹⁶⁸ Henle 2001: S.67-69.

¹⁶⁹ Götze: 05.02.2022, siehe Anhang 1.

¹⁷⁰ Henle 2001: S.65-67.

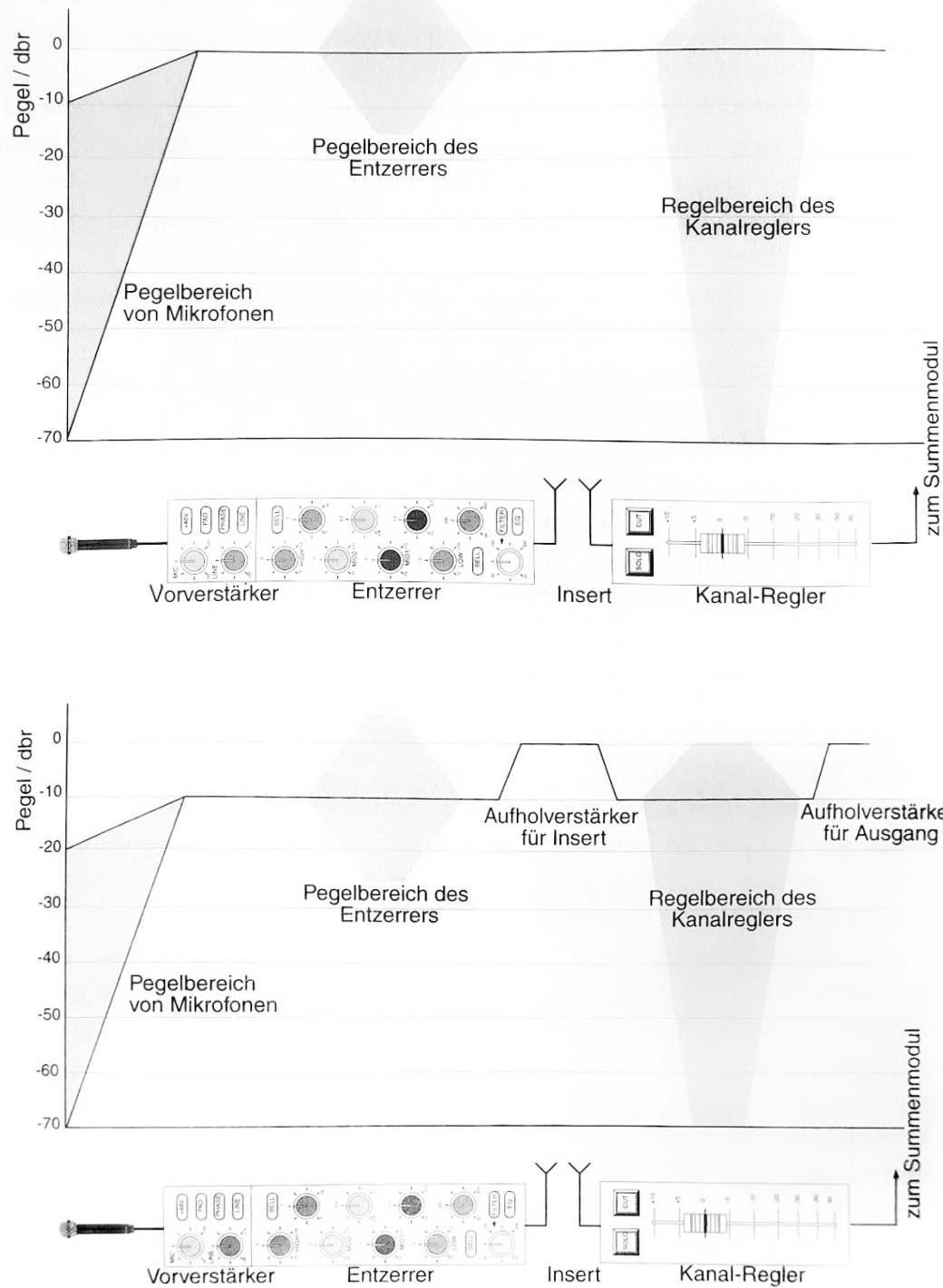


Abbildung 27 Pegeldiagramm einer hochwertigen Übertragungskette (oben) und einer Übertragungskette mit geringem internem Pegel (unten) ¹⁷¹

¹⁷¹ Henle 2001: S.66-67.

5.1 Vorverstärkung

Nachdem die vom Sprecher abgestrahlte Sprache vom Mikrofon in elektrische Spannung umgewandelt wurde, bedarf es weiterer Maßnahmen. Bevor die Sprache aufgenommen wird, muss der Signalpegel des Mikrofons vorverstärkt werden.¹⁷² Wird der Mikrofonvorverstärker falsch eingestellt, kommt es zu nichtlinearen Verzerrungen, die mithilfe eines Klirrfaktors eingeschätzt werden können. Das verstärkte Sprachsignal beinhaltet Störsignale und verfügt über einen bestimmten Dynamikbereich. In einer analogen Signalbearbeitung vor dem Aufnahmemedium kann die Dynamik der aufzunehmenden Sprache mithilfe eines Kompressors vorab eingeschränkt werden.

5.1.1 Mikrofonvorverstärker

Mikrofone erzeugen bei lauter Beschallung einen Spannungspegel zwischen -50 dBu und -30 dBu.¹⁷³ Das für folgende Geräte vorliegende Mikrofonsignal liegt zunächst als sehr geringe Wechselspannung vor. Demnach sind der Stromfluss und die übertragene Leistung vom Mikrofon sehr gering. Die niedrigen Pegel der Mikrofone müssen mit einem Vorverstärker (Pre Amp) angehoben werden.¹⁷⁴ Wie bei hochwertigen Mikrofonen, sollten Vorverstärker im Tonstudio einen möglichst linearen Übertragungsbereich haben.¹⁷⁵

Physikalisch bedingte Innenwiderstände der Mikrofone unterscheiden sich je nach Modell. Damit Geräte unmittelbar hintereinandergeschaltet werden können, führen Vorverstärker eine Impedanzwandlung durch. Das liegt daran, dass sich Spannungen im Verhältnis der Widerstandswerte von Geräten (Geräteimpedanzen) verteilen. Für die Spannungsübertragung sind Eingangs- und Ausgangswiderstände von elementarer Bedeutung. Bei der Spannungsübertragung fällt Verlustspannung am Ausgangswiderstand vom Quellgerät ab. Die Eingangsimpedanz des Zielgerätes muss mindestens fünfmal so groß wie die Ausgangsimpedanz des Quellgerätes sein (Überanpassung, Spannungsanpassung). Für eine optimale Leistungsübertragung müssen beide Impedanzen gleich hoch sein. Früher war diese sogenannte Leistungsanpassung üblich, bei der die Impedanzen von Quell- und Zielgerät auf 600 Ohm festgelegt waren. Heute wird die Leistungsanpassung für den Zeitcode (Timecode) verwendet. Für Studiopegel ordnet sich die Impedanz der Geräteeingänge bei oberhalb von 10 kHz ein. Die Ausgangsimpedanz liegt bei wenigen hundert Ohm. Je nach Wandlertyp heben Verstärker auf eine höhere Spannung von 0,775 V (Normalpegel) an oder nehmen eine Spannungsveränderung für einen nachgeschalteten Wandler vor. Lineare Verzerrungen der elektroakustischen Wandler können durch einen Vorverstärker korrigiert werden.¹⁷⁶

¹⁷² Stotz 2019: S.88.

¹⁷³ Warstat, Görne 2002: S.31.

¹⁷⁴ Webers 2007: S.140.

¹⁷⁵ Smyrek 2012: S.149.

¹⁷⁶ Warstat, Görne 2002: S.31-32.

Der Spannungspegel, den Studiogeräte bei Vollaussteuerung am Eingang erwarten und am Ausgang ausgeben, ist der Studiopegel oder Leitungspegel (Line-Pegel). Die Vollaussteuerung ist am Mischpult oder in der DAW bei 0 dB gekennzeichnet. Je nach Hersteller ist ein Pegel von +4 dB (1,23 V) oder +6 dB (1,55 V) möglich. Bei der Verschaltung verschiedener Geräte machen unterschiedliche Aussteuerungen um 2 dB keinen nennenswerten Unterschied. Pegelprobleme können bei der Verschaltung von professionellen und semiprofessionellen Geräten auftreten. Grund dafür ist, dass semiprofessionelle Geräte mit einer Spannung von 0,32 V (-10 dB) arbeiten. Einige professionelle Geräte können durch eigenbaute Schaltungen trotzdem in einer solchen Kombination verwendet werden.¹⁷⁷

Der Mikrofonpegel wird auf einen Normalpegel von +6 dB angehoben, wodurch Dämpfungsverluste durch eine Regieschaltung kompensiert werden.¹⁷⁸ Die Signalstärke wird dabei mit Reglern manipuliert. Studioverstärker unterliegen hohen Anforderungen hinsichtlich geringer Verzerrungen im Übertragungsbereich. Ein großer Übertragungsbereich und ein hoher Fremdspannungsabstand ist erwünscht. Des Weiteren sollten Studioverstärker betriebssicher, robust, langlebig und unempfindlich gegenüber elektrischen und magnetischen Verzerrungen sein.

Trotz einer Vollaussteuerung von +6 dBm (1,55 V) am Verstärkerausgang kann es sein, dass der Verstärker höher angesteuert wird. Man hat sich deshalb auf eine maximale Aussteuerung eines Studioverstärkers auf +22 dBm festgelegt. Damit äußere Einflüsse auf die Regieschaltung geringgehalten werden, wird ein Mikrofonpegel auf einen Normpegel von +6 dB in einem Vorverstärker angehoben.

Der Mikrofonverstärker, der sich für Sprachaufnahmen eignet, ist der Portico 511 von Rupert Neve Designs. In 6 dB-Schritten lässt sich die Mikrofonverstärkung bis zu 66 dB einstellen. Mit dem Regler Trim kann die Verstärkung stufenlos zwischen -6 dB und +6 dB feinjustiert werden. Eine im Verstärker verbaute LED-Pegelanzeige zeigt den Signalpegel an. Viele Vorverstärker sind mit einem Hochpassfilter ausgestattet, der tieffrequenten Trittschall aus dem Mikrofonsignal filtert. Per Potentiometer kann man den Hochpass zwischen 20 Hz und 250 Hz einstellen. Eine Phantomspeisung von 48 V für Kondensatormikrofone wird am Verstärker eingeschaltet und die Phase des Mikrofonsignals kann umgekehrt werden. Mit dem Regler Texture wird bei eingeschaltetem Silk-Knopf eine zusätzliche Schaltung in die Vorverstärkung integriert. Der Schaltkreis sorgt für eine vintageartige Wärme im Audiosignal.¹⁷⁹

Bei Sprachaufnahmen ist zu beachten, dass nur aus einer richtigen Aussteuerung aller Geräte im Übertragungsweg eine hochwertigen Aufnahme hervorgeht. Ist die Aussteuerung des Verstärkers zu niedrig, wird das System untersteuert was mit einem zu hohen Geräuschpegel einhergeht. Bei einer Übersteuerung treten nichtlineare Verzerrungen auf.¹⁸⁰

¹⁷⁷ Warstat, Görne 2002: S.31-32.

¹⁷⁸ Webers 2007: S.298.

¹⁷⁹ Rupert Neve Designs: 511, in: Rupert Neve Designs, 2022 [online] <https://rupertneve.com/products/511> [10.01.2022]

¹⁸⁰ Webers 2007: S.141.

5.1.2 Nichtlineare Verzerrungen und Klirrfaktor

Jedes Übertragungssystem verfügt über eine Aussteuerungsgrenze. Ein Eingangssignal, das für den am Verstärker eingestellten Arbeitspegel zu groß ist, beeinflusst die Signalverstärkung maßgeblich. Es unterliegt einer nichtlinearen Verzerrung. Zwar vermindern drastische nichtlineare Verzerrungen die Sprachverständlichkeit nur in begrenztem Maße, sie verfälschen aber das akustische Sprachsignal.¹⁸¹ Die Folge einer solchen Übersteuerung ist eine klangliche Verfärbung (Klirrvverzerrung). Nichtlineare Verzerrungen können, anders als lineare Verzerrungen nicht rückgängig gemacht werden. Deshalb gilt es diese bei Sprachaufnahmen für Hörbücher im Tonstudio zu vermeiden.

Laut der Norm DIN 45403 sind nichtlineare Verzerrungen Veränderungen der Kurvenform von Vorgängen eines Übertragungssystems durch nicht-lineare Glieder. Besitzt ein Übertragungssystem wie ein Verstärker eine nicht-lineare Kennlinie, unterliegen Signale nichtlinearer Verzerrungen. Die Kennlinie ist die amplitudenabhängige Beziehung zwischen Ein- und Ausgangssignal einer Schwingung, die Schwingungsformen nicht beeinflusst. Eine lineare Kennlinie lässt sich als Gerade darstellen. Eine Ausgangsspannung U_A wird dabei jedem Wert der Eingangsspannungen U_E zugeordnet. Die Steigung der Gerade ergibt sich aus dem Faktor, wie viel größer oder kleiner ein bestimmter Wert von U_A im Vergleich zum zugehörigen Wert von U_E ist. Bei nichtlinearen Kennlinien wird die Kurvenform so verändert, dass das Signal verzerrt wird. Die verzerrte Kennlinie weist auf den häufig auftretenden Fall der Klirrvverzerrung oder Begrenzung (Clipping) hin. Signale werden bei linearen Verzerrungen frequenzabhängig verformt, jedoch hat die Frequenz eines Signals keine Auswirkung bei nichtlinearen Verzerrungen.¹⁸²

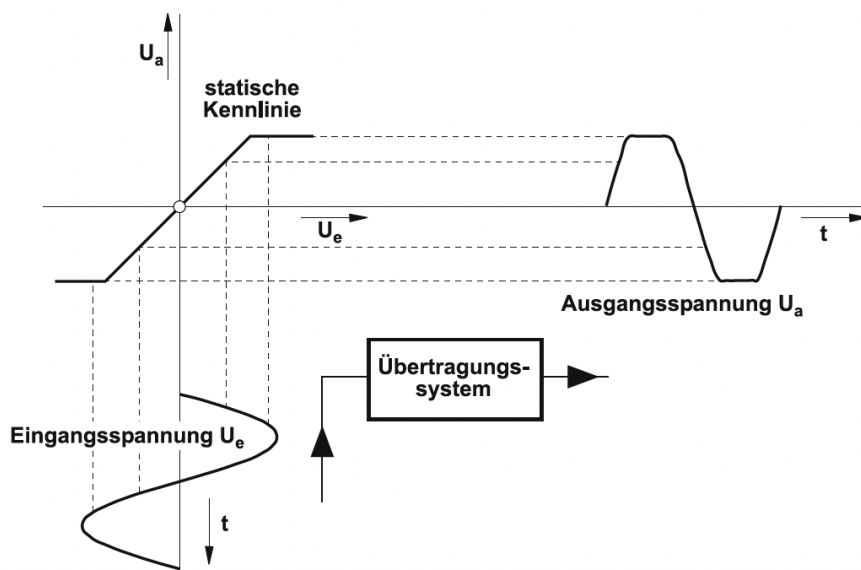


Abbildung 28 Eine Sinusspannung wird durch eine verzerrte Kennlinie verformt¹⁸³

¹⁸¹ Ernst 1998: S.6.

¹⁸² Henle 2001: S.414-416.

¹⁸³ Stotz 2019: S.6.

Klanglich lassen sich nichtlineare Verzerrungen daran erkennen, dass das verzerrte Signal schriller als das nicht verzerrte Signal klingt. Der Grund dafür ist, dass sich beim Clipping Spektralanteile zum Signal addieren, die im Eingangssignal nicht enthalten sind. Wenn ein System übersteuert wird, liegt eine harmonische Verzerrung vor. Es addieren sich Obertöne, die höher als das nicht verzerrte Signal sind, zum ursprünglichen Eingangssignal. Die addierten Harmonischen sind bei reinen Schwingungen deutlich zu hören.¹⁸⁴ Viele Signale bestehen allerdings nicht nur aus einem Ton mit klar erkennbarer Grundfrequenz. Wenn nur wenige Obertöne addiert werden, ist der Unterschied umso schwerer zu erkennen. Zudem werden hochfrequente Anteile des verformten Signals schwächer vom menschlichen Gehör wahrgenommen. Deshalb kann physiologisch die hörbare Abweichung oft nur mühsam ermittelt werden.

Verlässt eine verformte Sinuskurve ein Übertragungssystem mit nichtlinearen Verzerrungen, addieren sich zum reinen Sinussignal Obertöne. Da jede Schwingung als Summe verschiedener Sinusschwingungen dargestellt werden kann, lässt sich das auf alle Schwingungsverläufe übertragen. Neben nicht-linearen Kennlinien gibt es auch quadratische und kubische Kennlinien, die von linearen Kennlinien überlagert werden. Abhängig von der Kurvenform entstehen verschiedene Verzerrungen. Aus einer kubischen Kennlinie resultiert beim Clipping eine symmetrische Begrenzung. Es addieren sich Harmonische in ungeraden ganzzahligen Verhältnissen zur Grundfrequenz. Bei quadratischen Kennlinien treten unsymmetrische Verzerrungen auf. Gerade Vielfache der Frequenz des nicht verzerrten Signals addieren sich zur Schwingung. Harmonische weisen je nach Schwingungsform verschiedene Verteilungen der Vielfachen auf.¹⁸⁵ Quadratische Verzerrung sind mit untergestellter 2 und kubische Verzerrungen mit untergestellter 3 gekennzeichnet. Neben regelmäßigen Obertonverteilungen sind auch beliebige Konstellationen möglich. Zum Beispiel ist für symmetrische Begrenzungen eine Angabe der geraden Harmonischen nicht vorgesehen.¹⁸⁶

Formel 12 Klirrfaktor k

$$k = \sqrt{\frac{U_2^2 + U_3^2}{U_1^2 + U_2^2 + U_3^2 + \dots}}$$

¹⁸⁴ Görne 2002: S.175-177.

¹⁸⁵ Henle 2001: S.414-416.

¹⁸⁶ Warstat, Görne 2002: S. 25-27.

Mit dem Klirrfaktor k wird die Intensität dieser Obertöne in Prozent angegeben. Er berechnet sich aus der Wurzel des Effektivwertverhältnisses aller Obertöne (Klirrkompente) zur Gesamtspannung.¹⁸⁷ Die Ordnung der Harmonischen wird durch Indizes festgelegt. Der Gesamtklirrfaktor (Total Harmonic Distortion, THD) ist das Verhältnis zwischen Schwingungen, die durch eine nichtlineare Verzerrung entstanden sind, und dem verzerrten Gesamtsignal. Er errechnet sich aus den Effektivwerten der Spannung und bezieht sich je nach Schwingungsform auf bestimmte Harmonische. Eine Klirrdämpfung wird in manchen Fällen auch für Begrenzungen angegeben.¹⁸⁸

Formel 13 Klirrdämpfung a_k

$$a_k = 20 \cdot \log \frac{1}{k} = -20 \cdot \log k \quad dB$$

Verschiedene automatisierte Realisierungen bauen auf dem Vorgehen auf. Das Grundprinzip zur Ermittlung des Klirrfaktors funktioniert wie folgt. Zuerst wird der Effektivwert des Gesamtsignals gemessen und das Ergebnis auf 100% deklariert. Anschließend wird mit einem Filter die Grundfrequenz des gemessenen Signals unterdrückt. Alle Obertöne werden dadurch herausgefiltert. Jetzt kann der Klirrfaktor abgelesen werden. k ist in der Regel eine von der Frequenz f abhängige Funktion. Sie steigt mit der Frequenz. Deshalb ist die Spezifikation der Messfrequenz k . Ein weiteres Verfahren setzt mit einem Rauschimpuls ein. Ein Algorithmus prüft auf Korrelationen und wertet den Impuls aus, woraus der Klirrfaktor resultiert.¹⁸⁹

Bei einem falsch eingestellten Arbeitspegel am Verstärker ist das Eingangssignal nach der Verstärkung verzerrt. Aus diesem Grund ist es wichtig bei Sprachaufnahmen einen angemessenen Arbeitspegel am Verstärker einzustellen. Dieser muss so gewählt werden, dass der maximal zulässige Signalpegel zu keinem Zeitpunkt während der Aufnahme überschritten wird. Ein Bereich vom maximal zugelassenen Pegel bis einige dB darunter (Headroom) dient als Puffer bei überdurchschnittlich hohen Eingangspegeln.¹⁹⁰ In der Praxis erweist sich ein Pegel zwischen -18 dB und -12 dB am Vorverstärker bei durchschnittlich dynamischer Sprache als geeignet. Während der gesamten Aufnahme ist zu beachten, dass das Sprachsignal als Eingangssignal nicht übermäßig vom eingestellten Pegel abweicht. Bei einer mehrstündigen Hörbuchaufnahme kann die durchschnittliche Sprechlautstärke schwanken. Gegebenenfalls muss der Pegel angepasst werden, oder der Sprecher auf die sich veränderte durchschnittliche Lautstärke seiner Sprache aufmerksam gemacht werden.

¹⁸⁷ Stotz 2019: S.7.

¹⁸⁸ Smyrek 2012: S.101-103.

¹⁸⁹ Stotz 2019: S.8.

¹⁹⁰ Warstat, Görne 2002: S.115.

5.1.3 Störabstand und Dynamikbereich

Bei Aufnahmen verschlechtert sich die Übertragungsqualität durch Störsignale. Störsignale sind unabhängig vom Eingangssignal, das in ein Übertragungssystem hineingeht (Nutzsignal). In der Praxis addiert sich bei Sprachaufnahmen für Hörbücher stets Rauschen zur aufgenommenen Sprache. Störsignale entstehen durch Molekülbewegungen in elektrischen Bauteilen, die ohne angelegte Spannung Rauschen erzeugen. Verschiedene Bestandteile von Störsignalen werden spezifiziert und können messtechnisch erfasst werden.¹⁹¹ Der Nutzsignalpegel wird angegeben, um die zu ihm in Bezug stehende Störgröße zu ermitteln. Ohne einen angegebenen Gesamtsignalpegel kann der Störabstand nicht sinnvoll eingeordnet werden. Der in der Regel etwas höhere Gesamtsignalpegel entspricht idealerweise dem Nutzsignalpegel. Er ist die Summe aus Nutzsignal und Störsignal.

Durch Rauschen verursachte Störungen werden von dem Geräuschspannungsabstand berücksichtigt. Der Geräuschspannungsabstand ist die Pegeldifferenz zwischen dem Vollaussteuerungspegel und einer weiteren Störgröße, dem bewerteten Fremdpegel.¹⁹² Er wurde auf 0 dB festgelegt. Ohne die Angabe dieser Größe, kann das Störgeräusch gemessen werden. Der in dB angegebene Störgeräuschpegel liegt unterhalb des Gesamtsignalpegels, wenn er mit dem Geräuschspannungsabstand angegeben wird. Der Geräuschspannungsabstand ist das 20-fach logarithmierte Verhältnis aus der Gesamtspannung und der Störgeräuschspannung. Es erweist sich als sinnvoll das Störsignal über einen Filter zu schicken, da die Empfindlichkeit des menschlichen Ohres unterschiedliche Frequenzgänge für verschieden laute Schallereignisse aufweist (Kurven gleicher Lautstärke). Deswegen werden 30-Phon-Ohrkurvenfilter, sogenannte dB(A)-Filter, bei der Messung des bewerteten Geräuschspannungsabstandes (weighted) eingesetzt. Der Bezug auf die 30-Phon-Kurve wird dadurch gerechtfertigt, dass abgestrahlte Schallereignisse durchschnittlich mit etwa 30 Phon gehört werden.¹⁹³ Dabei darf nur ein Filter zwischengeschaltet werden und der bei der Messung verwendete Filter muss angegeben werden.

Ein Filter für die Messung des Fremdspannungsabstandes ist nicht vorgesehen. Als Fremdspannungsabstand wird der Störabstand zu allen Störsignalen bezeichnet. Er berechnet sich aus dem 20-fach logarithmierten Verhältnis aus der Gesamtspannung und der Fremdspannung und wird in dB angegeben. Zahlenmäßig fällt der Fremdspannungsabstand kleiner als der Geräuschspannungsabstand aus. Dabei werden beide Größen nach der deutschen Norm DIN 45405 oder dem internationalen Pendant CCIR 468 angegeben.¹⁹⁴

¹⁹¹ Henle 2001: S.419.

¹⁹² Webers 2007: S.713.

¹⁹³ Webers 2007: S.713.

¹⁹⁴ Henle 2001: S.420-422.

Weil der Geräuschspannungsabstand nützlichere Aussagen zur physiologischen Wirksamkeit gibt, wird er häufig statt dem Fremdspannungsabstand angegeben. Angaben zum Fremdspannungsabstand findet man insbesondere bei Geräten, die zu Brumm- und Rumpelgeräuschen neigen.

Der nutzbare Dynamikbereich gibt an, wie groß der Pegelanstand zwischen den leisesten und lautesten Signalen ist, die vom System gerade noch so verarbeitet werden können. Störsignale beeinflussen die untere Pegelgrenze, während die Übersteuerungsgrenze die obere Pegelgrenze bestimmt. Ein Sicherheitsabstand von etwa 6 dB sollte zum Störsignal vorhanden sein. Ansonsten hebt sich das Nutzsignal nicht ausreichend vom Störsignal ab und der Störpegel überdeckt das Nutzsignal. Geht man von dieser Annahme aus, muss die Dynamik in jedem Fall kleiner als der Fremdspannungsabstand sein.¹⁹⁵

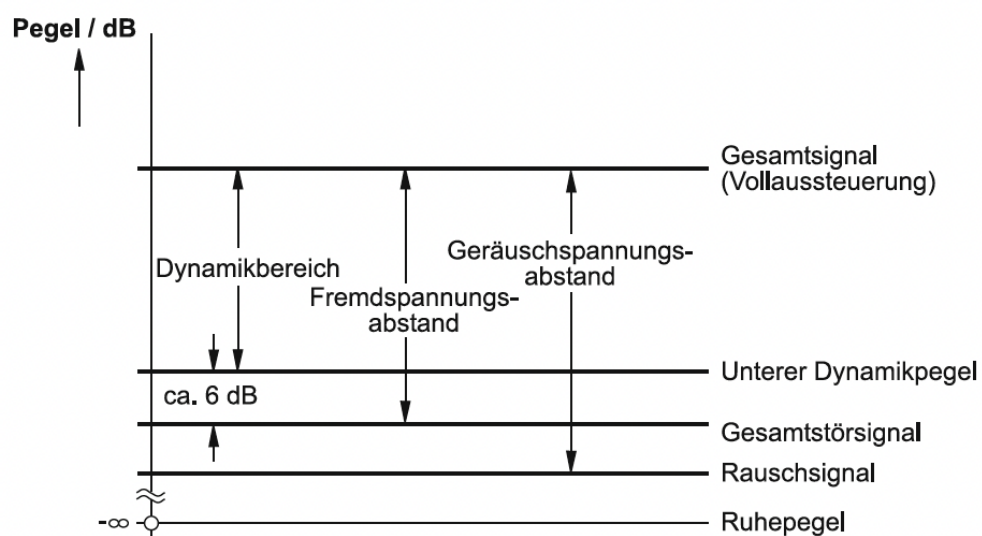


Abbildung 29 Störabstände im Übertragungssystem¹⁹⁶

In der Praxis findet man in vielen Fällen die englische Angabe für Störsignale Signal-to-Noise Ratio. Anders als die spezifizierten Störsignale gibt sie oft keine Auskunft darüber ob bewertet oder unbewertet gemessen wurde. Häufig geht aus der Größe weder der Bezugspegel hervor noch ob der Geräuschspannungsabstand oder der Fremdspannungsabstand gemeint ist. Deshalb ist die Angabe für einen Vergleich mit anderen Angaben selten verlässlich.¹⁹⁷

Der Anteil an Störsignalen ist bei Hörbuchaufnahmen möglichst gering zu halten. Bei der Auswahl der Aufnahmetechnik ist zu beachten, dass die angegebenen Störgrößen angemessen miteinander verglichen werden können. Gegebenenfalls beziehen sich Größen auf unterschiedliche Werte, was die Abwägung erschwert. Ein Überblick über die einzelnen Größen ermöglicht differenziertes Beurteilen von Angaben unterschiedlicher Hersteller.

¹⁹⁵ Stotz 2019: S.11-13.

¹⁹⁶ Stotz 2019: S. 13.

¹⁹⁷ Henle 2001: S.420.

5.2 Signalbearbeitung

Bei der analogen Bearbeitung, während der Sprachaufnahme muss bedacht vorgegangen werden. Alle Einstellungen an den analogen Geräten zur Echtzeitbeeinflussung können nachträglich nicht bearbeitet werden. Zu starke Bearbeitung kann die Qualität des Hörbuchs im ganzen folgenden Produktionsablauf beeinträchtigen. Deswegen bestimmt eine angemessene Einstellung der signalbearbeitenden Geräte das Aufnahmeergebnis maßgeblich. Auf analoge Signalbearbeitung bei der Aufnahme von Sprache für Hörbücher kann verzichtet werden. Allerdings bietet sie bei korrekter Anwendung Vorteile.

5.2.1 Kompression

Ein analoger Kompressor wird bei Aufnahmen für Hörbücher verwendet, um die Originaldynamik der Sprache anzupassen. Die Dynamik gibt Auskunft über die Pegeldifferenz zwischen lautestem und leisestem Signal, das von einer Schallquelle abgestrahlt wird. Als Regelverstärker ändert ein Kompressor die Dynamik der aufgenommenen Sprache, indem er sie einschränkt. Die Dynamik der Sprache als Schallereignis ist so groß, dass die Intentionen der Sprecher bei der Verarbeitung und Aufnahme des Sprachsignals ohne Kompression nicht erhalten bleibt. Nach korrektem Einschränken der Originaldynamik des Sprachsignals, erscheint die Sprache subjektiv lauter und gleichmäßiger. Folglich wirkt komprimierte Sprache verständlicher, voller und präsenter verglichen mit nicht komprimierter Sprache. Grund dafür ist die Verstärkung leiser Signale und Dämpfung lauter Signale durch den Kompressor. So kann bei Hörbuchaufnahmen im Tonstudio die Sprache mittels analoger Echtzeitbearbeitung verdichtet werden.

Es ist ratsam bei der Aufnahme für eine Hörbuchproduktion das Sprachsignal analog zu komprimieren, um einzelne große Pegelspitzen abzufangen. Das leicht komprimierte Signal wird anschließend verdichtet gewandelt. Da Sprache naturgemäß Schallereignisse mit sehr großem Dynamikumfang einschließt, ist es ratsam die Kompression nicht übermäßig einzusetzen. Besonders emotionale Passagen leben von Abweichungen vom durchschnittlichen Sprachpegel. Sprache, die zu stark komprimiert wurde, klingt unnatürlich, was für Hörbuchaufnahmen zu vermeiden ist.¹⁹⁸ Kompression wird auch zur Pegelbegrenzung eingesetzt, um einen bestimmten Pegel nicht zu überschreiten. So wird insbesondere bei Live-Aufnahmen Übersteuerung minimiert.

¹⁹⁸ Götze 04.02.2022, siehe Anhang 1.

Ein Kompressor kann in Reihe nach dem Vorverstärker und vor den Analog-Digital-Wandler (A/D-Wandler) geschaltet werden. Das Sprachsignal wird durch den Kompressor für alle folgenden Geräte verändert. Die Bearbeitung kann nicht rückgängig gemacht werden. Das komprimierte Signal wird nach der Analog-Digital-Wandlung aufgenommen. Deshalb ist es von elementarer Bedeutung, dass alle Einstellungen am Kompressor vor der Sprachaufnahme richtig eingestellt wurden.

Es bedarf einiger Experimentierfreudigkeit, um richtige Einstellungen für unterschiedliche Sprecher zu finden. Um bei erneuten Sprachaufnahmen mit demselben Sprecher Zeit zu sparen, empfiehlt es sich daher mittels Fotos von den Einstellungen am analogen Gerät die Parameter zu archivieren. Hierfür ist es notwendig alle Parameter zu kennen, die die Kompression charakterisieren. Je nach künstlerisch erwünschter und aufnahmetechnisch sinnvoller Komprimierung sind einige Einstellungen am Kompressor vorzunehmen. Für Sprachaufnahmen ist eine weiche und ausgeglichene Kompression erwünscht. Die Funktionsweise eines Kompressors kann man sich, wie einen aufmerksamen Toningenieur, der bei Lautstärkeschwankungen den Aufnahmepegel stetig nachregelt, vorstellen. Relativ lange Reaktionszeiten und die beschränkte Ausdauer von Menschen im Vergleich zu Maschinen rechtfertigen, dass für diese Korrekturen technische Mittel zu einem verlässlicheren Ergebnis führen. Ein Kompressor ist ein Gerät, das dazu dient, den momentanen Signalpegel ständig anzupassen. Mithilfe einer Kennlinie wird der Verstärkungsverlauf eines Kompressors dargestellt. Sie zeigt den Pegel des Ausgangssignals in Abhängigkeit vom Pegel des Eingangssignals.

Mit dem Schwellwert (Threshold) wird der Eingangswert definiert bei dessen Unter- oder Überschreiten der Kompressor auf das Signal einwirkt. Der durch den Schwellwert festgelegte zulässige Wirkungsbereich des Kompressors wird anhand des Voltage Controlled Amplifiers (VCA) zurückgeregelt. Der Schwellwert wird für Hörbuchaufnahmen so eingestellt, dass der Kompressor nur bei einzelnen Spitzenwerten des Signals oder bei Lauten mit überdurchschnittlich hohem Pegel arbeitet.¹⁹⁹ Konventionelle Kompressoren funktionieren nach dem Prinzip des Schwellwerts. Eine weitere Gattung der Kompressoren setzt unabhängig vom Schwellwert ein. Bei dem Soft Knee- oder Overeasy-Verfahren steigt das Kompressionsverhältnis mit dem Pegel des Eingangssignals. Daraus ergibt sich ein wärmerer Klang und eine einfachere Handhabung im Vergleich zu konventionellen Kompressoren.

Das Kompressionsverhältnis (Compression Ratio) bestimmt wie weit das Signal angehoben oder abgesenkt wird. In Abhängigkeit von der Änderung des Eingangspegels ändert sich der Ausgangspegel in einem bestimmten Betrag. Dieser Betrag wird mit dem Kompressionsverhältnis R angegeben. Es berechnet sich aus dem Verhältnis von der Dynamik des Eingangspegels zur Dynamik des Ausgangspegels.

¹⁹⁹ Götze 04.02.2022, siehe Anhang 1.

Formel 14 Kompressionsverhältnis R

$$R = \frac{\Delta P_E [\text{dB}]}{\Delta P_A [\text{dB}]} = \tan \beta$$

Bei einem Kompressor ist das Kompressionsverhältnis mindestens eins. Beträgt es 2:1, bewirkt der Kompressor eine Eingangspegeländerung von 2 dB in eine Ausgangspegeländerung von 1 dB. Ist das Kompressionsverhältnis größer als 15:1, wird das Signal nahezu limitiert, wenn es den Schwellwert überschreitet.²⁰⁰ Nach anderen Quellen liegt das Kompressionsverhältnis von Kompressoren zwischen 1,8:1 und 8:1.²⁰¹ Nähert sich der Wert ∞ , gleicht die Wirkung des Kompressors der eines Begrenzers (Limiter). Einige Kompressoren mit hohen Kompressionsverhältnissen tragen daher zusätzlich die Bezeichnung Limiter. Ein Begrenzer ist ein Verstärker, bei dem die Ausgangsspannung proportional zur Eingangsspannung bis zu einem Maximalwert ist. Steigt die Eingangsspannung über den Maximalwert, begrenzt sie sich auf den maximalen Wert.²⁰²

Wie die analoge Kompression kann eine analoge Begrenzung von Signalen nicht rückgängig gemacht werden. Sprachsignale erfahren durch Limiter einen weitaus stärkeren Dynamikverlust als durch Kompressoren. Folglich bezwecken Begrenzer eine deutlichere Klangverfärbung als Kompressoren. Bei Sprachaufnahmen für Hörbücher wird ergo keine Begrenzung eingesetzt. Um das Sprachsignal in der Aufnahme zu verdichten, wählt man ein geringes Kompressionsverhältnis bis maximal 4:1.²⁰³ Eine Ausnahme bildet der Einsatz von Limitern zur Verhinderung von Übersteuerung. Dabei limitiert der Begrenzer das Sprachsignal nur in den Pegelspitzen. Der Limiter wird bei dieser Anwendung häufig hinter den Kompressor geschaltet. Bei korrekt eingepegelttem Signal, kann auf den Limiter verzichtet werden, was für geschulte Sprecher bei Hörbuchaufnahmen im Tonstudio der Fall ist.

Die Reaktionszeit des Kompressors wird durch den Parameter der Ansprechzeit (Attack oder Response) festgelegt. Die Größe bestimmt wann der Kompressor, nachdem das Signal den Schwellwert überschreitet, einsetzt. Überschreitet das momentane Signal den Schwellwert nicht mehr kehrt der Kompressor wieder zu seiner Normalstellung zurück. Das zuvor herunter- und heraufgeregelter Signal liegt dann im originalen Ausgangspegel vor. Der Regler für die Rückstellzeit des Kompressors (Release) ist das Gegenteil der Ansprechzeit. Mit dem zugehörigen Regler wird eingestellt, wie lange es braucht, bis der Kompressor nicht mehr reagiert. Die Reaktionszeiten des Kompressors hängen stark von der Anwendung ab. Stellt man die Reaktionszeiten zu kurz ein, setzt eine Modulation stehender Signalanteile durch überlagerte Pegelspitzen ein. Dieser hörbare Effekt wird auch „Pumpen“ genannt²⁰⁴ und als unnatürlich empfunden.²⁰⁵ Bei der Kompression von Sprache sollte er deshalb vermieden werden.

²⁰⁰ Webers 2007: S.315.

²⁰¹ Henle 2001: S.273.

²⁰² Webers 2007: S.314.

²⁰³ Henle 2001: S.276.

²⁰⁴ Webers 2007: S.320.

²⁰⁵ Götze 04.02.2022, siehe Anhang 1.

Braucht der Kompressor sehr lange, um wieder zu seiner Ausgangsposition zurückzukehren, werden leise Signale, die unmittelbar nach dem Kompressor auslösende Signale folgen, zusätzlich heruntergeregelt. In der Praxis erweisen sich eine sehr schnelle Attack Zeit bis zu 2 ms und eine lange Release Zeit zwischen 200 und 500ms für Sprachkompression als brauchbare Werte.²⁰⁶ Da die Rückstellzeit von den unregelmäßigen Nachklingvorgängen des Sprachsignals abhängt, empfiehlt es sich eine Rückstellzeit zu wählen, die für die langen Ausklingvorgänge angemessen ist.

Tieffrequente Signale können durch einen unzureichend langen Rückregelvorgang zu einem erhöhten Klirrfaktor führen. Grund dafür ist, dass jede Teilwelle einer Schwingung anstatt die komplette Schwingung, begrenzt wird. Dadurch können Halbwellen erneut vom Kompressor oder Limiter angesprochen werden bevor der Rückstellvorgang beendet wurde. Das Problem wird umgangen, wenn die Begrenzungsphase verlängert wird. Der Rückstellvorgang wird hierfür mit einem Haltekreis verzögert. Daraus ergibt sich die sogenannte Holding Phase. So werden bereits angesprochene Halbwellen nicht doppelt bearbeitet. Der Regler Hold muss hierfür auf etwa 40 - 50 ms gestellt werden, damit Verzerrungen in den tiefen Frequenzen vermieden werden. Die Holding Phase lässt sich für Sprachaufnahmen mit Sprecherinnen verkürzen, wenn weniger tiefe Frequenzen vorliegen.

In vielen Fällen stellt eine optische Anzeige den momentan reduzierten Pegel auf dem Gerät an. Häufig handelt es sich um eine LED-Anzeige oder ein VU-Meter. Die Verstärkungsverminderung (Gain Reduction) wird durch Threshold und Ratio bestimmt. Über die Anzeige können Rückschlüsse auf die Wirkung des Kompressors gezogen werden. Wenn die Gain Reduction bei sehr niedrigen Eingangspegeln nicht mehr 0 dB erreicht, wurde beispielsweise der Threshold zu niedrig gesetzt.²⁰⁷ Weil sich der Lautstärkeindruck bei der Kompression verändert, bedarf es häufig einer Korrektur des komprimierten Signalpegels. Deshalb lässt sich mit den Regler Output Gain der absolute Ausgangspegel für nachfolgende Geräte anpassen.

Die Art und Weise wie der Pegel des zu komprimierenden Signals bestimmt wird, hat Einfluss auf den Klang und die Effektivität des Kompressors. Ein vor der Kompression gefiltertes Eingangssignals bewirkt, dass nur bestimmte Signalanteile komprimiert werden. Für diesen Zweck verfügen einige Kompressoren über einen zweiten Eingang (Side Chain). Das Eingangssignal geht gefiltert über den zweiten Eingang in den Kompressor und gleichzeitig direkt in den Kompressor. Mit dem gefilterten Signal wird bestimmt welche Frequenzen komprimiert werden sollen. Aus dieser Schaltung resultiert ein in bestimmten Frequenzbereichen komprimiertes Signal.

²⁰⁶ Götze 04.02.2022, siehe Anhang 1.

²⁰⁷ Henle 2011: S.274.

Der Lautstärkeindruck ist neben dem Pegel von der Frequenz eines Signals abhängig. Aus diesem Grund berücksichtigen die gefilterten Frequenzanteile im zweiten Eingangskanal die Höreigenschaften des menschlichen Gehörs. Hohe Frequenzbereiche werden bei komprimierten Signalen ohne gefiltertes Side Chain Signal lauter als tiefe Frequenzen wahrgenommen. Bei Sprachaufnahmen soll vermieden werden, dass die hohen Frequenzen von S- und Zischlauten hervortreten. Diese treten ab etwa 6 kHz auf. Um diesen Effekt zu erzielen, können im zweiten Eingangskanal die Frequenzbereiche der S- und Zischlaute herausgefiltert oder angehoben werden. Dadurch werden die mit lautem Pegel behafteten störenden Frequenzen vom Kompressor heruntergeregelt. Die Arbeitsweise eines De-essers funktioniert nach diesem Prinzip. Bei leisen Passagen hebt der Kompressor den Signalpegel an. Demnach werden Störgeräusche wie Schmatzen und überdurchschnittlich laute Atmen angehoben. Zudem wird Rauschen durch Kompression verstärkt. Abhilfe schafft ein eingebautes Noise Gate, das Rauschen in Pausen ausblendet. Die Einstellungen des integrierten Gates sind entweder an die Kompressorparameter gekoppelt oder können als Voreinstellungen gespeichert werden. Das Gate soll so klein wie möglich eingestellt werden. Ansonsten werden sehr leise Laute, die zur Sprache gehören unterdrückt, was zu starkem Qualitätsverlust führt. Spezielle Algorithmen in hochwertigen Kompressoren verhindern dessen Einsatz in Signalpausen. Diese Geräte verfügen über einen besseren Signal-Rausch-Abstand als einfachere Kompressoren. Darüber hinaus kann bei hochwertigen Kompressoren das Pumpen bei tiefen Frequenzen vermindert werden. Die Frequenzbänder eines zerlegten Eingangssignales werden separat komprimiert. Anschließend werden die Bänder wieder zu einem Gesamtsignal zusammengefügt. Solche Kompressoren verfügen über eine spezielle Eingangsfilterbank. Für jedes Frequenzband müssen separat Einstellungen vorgenommen werden, was den Vorgang aufwendig gestaltet.²⁰⁸

Automatische Anpassungen von Einstellungen sind möglich, aber für Sprachaufnahmen nicht empfehlenswert. Über Algorithmen, die den Signalverlauf analysieren, werden Parameter nachjustiert. Um Arbeit zu sparen, können Einstellungen für bestimmte Sprecher notiert werden. Kommt der Sprecher erneut zur Aufnahme in das Tonstudio können die Parameter gleich eingestellt werden und mit großer Wahrscheinlichkeit sind nur kleine tagesabhängige Änderungen vorzunehmen. Verschiedene Kompressoren beeinflussen den Klang des Sprachsignals unterschiedlich. Eine Klangbeurteilung kann jedoch nicht generalisiert werden. In professionellen Tonstudios bemüht man sich für spezielle Anwendungsfälle den subjektiv am besten klingenden Kompressor auszuwählen. Ein für Sprachaufnahmen geeigneter analoger Kompressor mit Limiterfunktion ist der Rupert Neve Designs Portigo 543.²⁰⁹ Er verfügt über alle gängigen Parameter zur Konfiguration. Hochwertige analoge Geräte anderer Hersteller sind ebenfalls möglich.

²⁰⁸ Warstat, Görne 2002: S.243-248.

²⁰⁹ Rupert Neve Designs: 543, in: Rupert Neve Designs, 2022 [online] <https://rupertneve.com/products/543> [14.01.2022]

6. Fazit

Hochwertige Sprachaufnahmen gehen aus dem Einsatz geeigneter Aufnahmetechnik und dem Wissen wie diese zu bedienen ist, hervor. Wenn bei Hörbuchaufnahmen auftretende Phänomene in der Theorie verstanden wurden, kann eine angemessene Realisierung in der Praxis erfolgen. So zeigt die vorliegende Bachelorarbeit die Wechselwirkung zwischen theoretischem Wissen über Besonderheiten bei Sprachaufnahmen im Tonstudio und Sprachaufnahmen für Hörbücher in der Praxis. Aus den in der Arbeit untersuchten Themen geht hervor, dass ein umfassendes Wissen über zahlreiche audioteknische Phänomene notwendig ist, um hochwertige Sprachaufnahmen für Hörbücher umzusetzen. Darüber hinaus trägt eine ausführliche Kenntnis über auditive Grundlagen und Sprache als Schallereignis dazu bei, das Aufnahmeverfahren zu optimieren.

Es wurde erarbeitet, inwieweit grundlegende Kenntnisse über Schall als Pegel und die menschliche Spracherzeugung konstruktiv zu qualitätsvollen Sprachaufnahmen beitragen können. Anhand des Quelle-Filter-Modells und dem Formant-Synthesizer wurde der Mechanismus der Spracherzeugung beleuchtet. Eine Frequenzanalyse des Sprachsignals zeigte Eigenschaften des Sprachsignals auf, bevor Unterschiede zwischen Sprechern charakterisiert wurden. Anschließend wurden verschiedene Phänomene und technische Aspekte untersucht, die zu einer für Sprachaufnahmen angemessenen Mikrofonierung beitragen. Des Weiteren erläuterte die Arbeit elementare Größen, die vor allem den Einsatz eines Kondensatormikrofons mit einer Nierencharakteristik begründen. Im Kapitel Mikrofonierung wurden Maßnahmen gegen Störschall belegt und die Positionierung des zu verwendenden Mikrofons ausgearbeitet. Sodann beschäftigte sich die Arbeit mit einer geeigneten analogen Echtzeitbeeinflussung des Sprachsignals. Die bei der Vorverstärkung des Signals auftretenden Besonderheiten, wobei charakterisierende Größen erläutert wurden. Das Kapitel analoge Signalbearbeitung legte die Kompression während der Aufnahme dar und wie die auf die Aufnahme folgende Postproduktion vermindert werden kann.

Zusammenfassend zeigt die Bachelorarbeit auf, dass Schall als mechanische Schwingungen im elastischen Medium Luft von Menschen abgestrahlt wird. Der Schall wird in der Audiotechnik als das logarithmische Verhältnismaß Schalldruckpegel angegeben. Bei Sprache entsteht das akustische Signal durch den menschlichen Sprachmechanismus. Ein aus der Lunge entweichender Luftstrom erzeugt Laute, indem Stimmbänder angeregt werden und einzelne Elemente des Sprachmechanismus bestimmte Stellungen einnehmen oder Bewegungen ausführen. Die Filterung des Luftstroms kann durch das Quelle-Filter-Modell simuliert werden. Ein Formant-Synthesizer imitiert menschliche Sprache anhand charakteristischer Resonanzfrequenzen der Stimme. Eine Frequenzanalyse des Sprachsignals ergibt, dass in einem Frequenz-Zeit-Spektrogramm spezifische Frequenzbereiche in Zusammenhang mit Amplitude und Bandbreite Laute identifiziert werden können. Trotz durchschnittlicher Werte unterscheiden sich Stimmen, wobei trainierte Sprecher in der Regel über längere Zeit eine gleichmäßige und deutlichere Stimme als Laiensprecher aufweisen.

Hochwertige Hörbuchaufnahmen bedürfen einer angemessenen Mikrofonierung. Schall wird von der Mikrofonkapsel in mechanische Schwingungen und anschließend durch die elektromechanische Wandlung in ein elektrisches Signal umgesetzt. Für Hörbuchaufnahmen wird ein Druckgradientenempfänger eingesetzt, der über die Richtcharakteristik Niere verfügt. Der Nahbesprechungseffekt tritt nur bei Gradientenempfängern auf. Er hebt bei geringen Aufnahmeabständen tiefe Frequenzen an und wird bei Sprachaufnahmen als künstlerisches Mittel eingesetzt. Mikrofone mit umschaltbarer Richtcharakteristik wie das OC818 von Austrian Audio bieten eine Möglichkeit den Klang einer Aufnahme zu beeinflussen. Die elektromechanische Wandlung geschieht bei Hörbuchaufnahmen in der Regel elektrostatisch mit einem Kondensatormikrofon mit einer Phantomspeisung von 48 V. Die Funktionsweise von Kondensatormikrofonen beruht auf der Kapazitätsänderung des im Wandler verbauten Kondensators. Der Frequenzgang und der Übertragungsbereich ist für die Mikrofonauswahl von klanglicher Bedeutung, wobei auffällt, dass Kondensator-Gradientenempfängern sich auch diesbezüglich für Hörbuchaufnahmen eignen. Weitere für die Aufnahme zu vergleichende Eigenschaften von Mikrofonen sind der Übertragungsfaktor und das Übertragungsmaß, die Nenn- und Nennabschlussimpedanz, der Störpegel und das Impulsverhalten. Kleinmembranmikrofone können die Stimme signalgetreuer umwandeln. Für Sprachaufnahmen bevorzugt man allerdings Großmembranmikrofone, weil eine große Mikrofonmembran durch ihre Trägheit nachschwingt. Störschall, der die Sprachaufnahmen negativ beeinflusst kann durch Maßnahmen wie Poppschutzvorrichtungen und federnde Mikrofonhalterungen und Ständer für Tritt- und Körperschall vermieden werden. Da die Stimme in Hörbuchproduktionen raumlos klingen sollte, ist das Mikrofon in einem Abstand von 20 bis 30 cm vor dem Sprecher zu platzieren. Oft hängt das Mikrofon an einer Halterung, sodass es über dem Lesertisch von leicht oben und seitlich auf den Mund des Sprechers gerichtet ist. Durch diese Position wird der Nahbesprechungseffekt optimal eingesetzt, der Kammfiltereffekt weitgehend vermieden und das Abstrahlverhalten des menschlichen Sprachapparates miteinbezogen.

Bei analoger Echtzeitbeeinflussung von Sprachsignalen bei Hörbuchaufnahmen bedarf es einer elektrischen Anpassung aller Geräte in der Übertragungskette. Die Arbeitspegel aller Geräte in der Übertragungskette müssen aufeinander abgestimmt sein. Das akustische Sprachsignal wird nach der elektromechanischen Wandlung vorverstärkt. Wird ein Übertragungssystem übersteuert, treten irreversible nichtlineare Verzerrungen auf. Diese können mit dem Klirrfaktor berechnet werden. Messtechnisch erfassbares Rauschen von Gerätebestandteilen verschlechtert die Übertragungsqualität bei Sprachaufnahmen. Bei der Angabe von Störsignalen ist auf die genaue Bezeichnung der Störgröße zu achten, um den Rauschanteil in einer Aufnahme wirksam beurteilen und gering halten zu können. Die Sprachaufnahme kann in Echtzeit analog bearbeitet werden. Ein analoger Kompressor dient dazu in geringem Maße die große Originaldynamik der Sprache anzupassen, sodass diese lauter und präsenter klingt. Da die analoge Signalbearbeitung nicht rückgängig gemacht werden kann, sollte bedacht vorgegangen werden.

6.1 Ausblick

Eine Alternative für hochwertige Hörbuchaufnahmen im Tonstudio sind Sprachaufnahmen über die Ferne (remote). Während meines Praktikums bei dem Tonstudio Tongoetze wurden zahlreiche Studioaufnahmen von zu Hause aus realisiert. Grund dafür war der beschränkte Studiobetrieb während der Corona Pandemie. Ziel einer Hörbuchaufnahme über remote ist es, dass der Sprecher zu Hause das Hörbuch einlesen kann. Während im Tonstudio oder Homeoffice tontechnische Aufgaben an der DAW und die Kontrolle des Gesprochenen übernommen werden, ist bei dieser Vorgehensweise der Sprecher für die Mikrofonierung und andere aufnahmetechnischen Praktiken verantwortlich. Die digitalen Audio Workstations von Sprecher und Toningenieur werden über die Online-Anwendung Session Link Pro verbunden.²¹⁰ Ein internes Routing von digitalen Interfaces mittels der Software Loopback²¹¹ auf dem Computer des Toningenieurs ermöglicht es das Mikrofonsignal des Sprechers über das Internet abzugreifen. Während der Aufnahme nehmen Sprecher und Toningenieur in ihrer DAW auf. Das liegt daran, dass die von der Internetqualität abhängenden Aufnahmen des Toningenieurs Artefakte enthalten. Während der Aufnahme erstellt der Toningenieur Notizen für die Postproduktion im DAW-Projekt. Im Anschluss an die Aufnahme übergibt der Sprecher seine lokal aufgenommenen Dateien an den Toningenieur. Dieser synchronisiert die lokalen Sprecherdateien mit der Aufnahme über Session Link Pro.

Nach den Sprachaufnahmen wird das Hörbuch geschnitten, editiert, korrigiert und gemastert. Die während den Aufnahmen entstandenen Notizen bieten einen wichtigen Anhaltspunkt für Schnitte, Pausenverlängerungen und inhaltlich bedingte Bearbeitungen. In der Postproduktion folgt eine digitale Nachbearbeitung. Plug-Ins, die eingesetzt werden, sind De-Esser, Gates, Equalizer, Kompressoren, Limiter und Metering Tools. Die Räumlichkeit der Sprachaufnahme kann mit dem Plug-In Voice De-noise von iZotope RX²¹² zusätzlich reduziert werden. Übermäßig laute Atemgeräusche können durch ein angesprochenes Gate (Side Chained Gate) eliminiert werden. Ein sprachoptimiertes Duplikat der Sprachspur spricht dabei ein Gate an, welches auf der originalen Spur der Aufnahme liegt. Die duplizierte Spur der Sprecheraufnahme wird so bearbeitet, dass möglichst nur Sprache zu hören ist. Nach Schnitt und Bearbeitung (Editing) werden alle Kapitel des Hörbuchs mit dem Verlag geteilt. Es folgen mehrere Korrekturschleifen. In der Regel dient eine weitere Aufnahme dazu, technische und sprecherseitige Fehler für die finale Fassung zu beheben. Nach der Postproduktion und den Korrekturen wird das Hörbuch gemastert veröffentlicht.

²¹⁰ SessionLinkPro Solutions GmbH: HOME, in: SessionLinkPRO, 2014-2022 [online] <https://de.sessionlinkpro.com/> [14.02.2022]

²¹¹ Rogue Amoeba Software: LOOPBACK. CABLE-FREE AUDIO ROUTING FOR MAC, in: ROGUE AMOBEA, 08.12.2021 [online] <https://rogueamoeba.com/loopback/> [14.02.2022]

²¹² iZotope: RX 9 Voice De-noise, in: iZotope, 2001-2021 [online] <https://www.izotope.com/en/products/rx/features/voice-de-noise.html> [16.02.2022]

Literaturverzeichnis

Bücher

Bernstein, Herbert: **Elektroakustik. Mikrofone, Klangstufen, Verstärker, Filterschaltungen und Lautsprecher**, 2. Auflage, Wiesbaden: Springer 2019.

Boré, Gerhart/ Peus, Peter: **Mikrophone. Arbeitsweise und Ausführungsbeispiele**, 5. Auflage, Berlin: Georg Neumann GmbH 2002.

Dickreiter, Michael: **Mikrofon-Aufnahmetechnik**, 3. Auflage, Stuttgart: Hirzel 2003.

Dickreiter, Michael: **Handbuch der Tonstudioteknik**, Bd. 1, Berlin: De Gruyter 2014.

Friesicke, Andreas: **Die Audio-Enzyklopädie. Ein Nachschlagwerk für Tontechniker**, 2. Auflage, Berlin: De Gruyter Saur 2014.

Görne, Thomas: **Mikrofone in Theorie und Praxis**, 6. Auflage, Aachen: Elektor 2002.

Henle, Hubert: **das Tonstudio Handbuch**, 5. Auflage, München: GC Carstensen 2001.

Hufschmid, Markus: **Grundlagen der Elektrotechnik. Einführung für Studierende der Ingenieur- und Naturwissenschaften**, Wiesbaden: Springer Vieweg 2021.

Pawera, Norbert: **Mikrofonpraxis. Tipps und Tricks für Bühne und Studio; Technik, Akustik und Aufnahmepraxis für Instrumente und Gesang**, 4. Auflage, München: Franzis 2003.

Raabe, Armin/ Holstein, Peter: **Akustik und Raumklima. Raumkomfortbewertung und Energieeffizienz**, Wiesbaden: Springer Fachmedien 2021.

Smyrek, Volker: **Tontechnik. Für Veranstaltungstechniker in Ausbildung und Praxis**, 2. Auflage, Stuttgart: Hirzel 2012.

Stotz, Dieter: **Computergestützte Audio- und Videotechnik. Multimediatechnik in der Anwendung**, 3. Auflage, Berlin: Springer 2019.

Terhardt, Ernst: **Akustische Kommunikation. Grundlagen mit Hörbeispielen**, Berlin/ Heidelberg: Springer 1998.

Warstat, Michael/ Görne, Thomas: **Studiotechnik Hintergrund und Praxiswissen**, 5. Auflage, Aachen: Elektor 2002.

Webers, Johannes: **Handbuch der Tonstudioteknik für Film Funk und Fernsehen**, 9. Auflage, Poing: Franzis 2007.

Weinzierl, Stefan: **Handbuch der Audiotechnik**, Berlin/ Heidelberg: Springer 2008.

Wirsum, Siegfried: **Mischpulte und Mischpultmodule. Tonquellen, Mischpulte und Bausteine – Anleitungen für den Selbstbau nach Maß**, 7. Auflage, München: Franzis 1988.

Internet

Bauer, Armin: **Test. Neumann U87 Ai, Grossmembran Mikrofon**, in: AMAZONA.de music can change the world, 2022 [online] <https://www.amazona.de/test-neumann-u-87-ai-grossmembran-mikrofon/3/> [07.02.22]

Dennis, L.: **Hören statt Lesen. Hörbücher, Hörspiele & Co. werden in Deutschland immer beliebter**, in: Forschung und Wissen, 22.05.2021 [online] <https://www.forschung-und-wissen.de/nachrichten/oekonomie/hoerbuecher-hoerspiele-und-co.-werden-in-deutschland-immer-beliebter-13374997> [05.11.2021]

Gräfe, Daniel: **Start von Bookbeat. Markt für Hörbücher wächst weiter**, in: Stuttgarter Zeitung, 23.07.2017 [online] <https://www.stuttgarter-zeitung.de/inhalt.start-von-bookbeat-markt-fuer-hoerbuecher-waechst-weiter.4be4c487-3054-4d3e-b105-cfd27f5a3031.html> [05.11.2021]

Hau, Andreas: **Austrian Audio OC818 Großmembran-Kondensatormikrofon im Test**, in: Sound & Recording, 11.09.2020 [online] <https://www.soundandrecording.de/equipment/austrian-audio-oc818-grossmembran-kondensatormikrofon-im-test/> [16.11.21]

HfMDK Frankfurt: **Töne, Klänge und Klangfarben**. Kapitel 1. Schwingungen und Wellen, in: Akustische Grundlagen, 2022 [online] https://www.selma.hfmdk-frankfurt.de/akustische-grundlagen/abb/tut_formant.png, [07.02.2022]

inVENTer GmbH: **Geräuscentwicklung und Schall bei der Lüftung. Grundlagen und Definition von Schall**, in: inVENTer, 2022 [online], https://www.inventer.de/wp-content/uploads/2021/03/csm_Schalldruck-und-Uebertragung-der-Molekuele_01d031e2f0.jpg [04.02.2022]

iZotope: **RX 9 Voice De-noise**, in: iZotope, 2001-2021 [online] <https://www.izotope.com/en/products/rx/features/voice-de-noise.html> [16.02.2022]

König & Meyer: **23860 Mikrofonarm**, in: König & Meyer. Stands for Music, 2022 [online] <https://www.k-m.de/produkte/mikrofonstative/mikrofonarme/23860-mikrofonarm-schwarz?c=128> [11.02.2022]

Maier, Peter/ Lembke, Stephan: **Raumakustik. Das Studio, der Eierkarton, die Absorber und Architektur**, in: SOUND & RECORDING, 2022 [online] <https://www.soundandrecording.de/tutorials/raumakustik-das-studio-der-eierkarton-die-absorber-und-architektur/> [08.02.22]

Rogue Amoeba Software: **LOOPBACK. CABLE-FREE AUDIO ROUTING FOR MAC**, in: ROGUE AMOEBEA, 08.12.2021 [online] <https://rogueamoeba.com/loopback/> [14.02.2022]

Rupert Neve Designs: **511**, in: Rupert Neve Designs, 2022 [online] <https://rupertneve.com/products/511> [10.01.2022]

Rupert Neve Designs: **543**, in: Rupert Neve Designs, 2022 [online]
<https://rupertneve.com/products/543> [14.01.2022]

SessionLinkPro Solutions GmbH: **HOME**, in: SessionLinkPRO, 2014-2022 [online]
<https://de.sessionlinkpro.com/> [14.02.2022]

Shure GmbH: **BETA58A. Gesangsmikrofon**, in: Shure Publications, 2022 [online]
<https://pubs.shure.com/guide/BETA58A/de-DE> [08.02.22]

STUDIOBOX GmbH: **STUDIOBOX Schallschutzkabine**, in: STUDIOBOX Schalldämmende Akustik Studios, 2022 [online] <https://www.studiobox.de/> [11.01.22]

Thomann GmbH: **Austrian Audio OC818 Studio Set**, in: Thomann, 2022 [online]
https://www.thomann.de/de/austrian_audio_oc818_studio_set.htm?gclid=CjwKCAiAo4OOBhBBEiwA5KWu_7vtNep0rzTJe1InmmO5uQ9Q9JDA82D7U9guAshaJluEiKQUWD0tLhoCVUEQAvD_BwE [07.02.22]

Thornton, Mike: **iZotope RX 9 Tested - The Good, The Bad And The Ugly**, in: PRO TOOLS EXPERT, 14.10.21 [online] <https://www.pro-tools-expert.com/production-expert-1/izotope-rx-9-tested-the-good-the-bad-and-the-ugly> [11.01.21]

Trommer, Jochen: **Resonanz. Das Quelle-Filter-Modell der Vokalproduktion**, in: Universität Leipzig Institut für Linguistik, 2007 [online] <https://www.uni-leipzig.de/>, [04.02.2022]

WERSÉNYI, György: **Grundlagen der Fahrzeugakustik**, in: UNIVERSITAS-GYÖR NONPROFIT Kft., 27.08.2014 [online]
<http://vip.tilb.sze.hu/~wersenyi/HTML/GRUNDLAGEN%20DER%20FAHRZEUGAKUSTIK.htm>, [04.02.2022]

Weyerer, Marlene: **Jetzt gibt's was auf die Ohren, denn: „Hören ist das neue Sehen“**, in: Augsburger Allgemeine, 12.06.2021 [online] <https://www.augsburger-allgemeine.de/kultur/Journal/Die-neue-Lust-am-Audio-Jetzt-gibts-was-auf-die-Ohren-denn-Hoeren-ist-das-neue-Sehen-id59862601.html> [05.11.21]

Anhang

1. Interview mit Markus Götze

Ein Interview mit dem Hörbuchproduzenten und Inhaber des Tonstudios Tongoetze
04.02.2022, Vereinfachte Transkription

Lucy [00:00:00] Hallo Markus, schön, dass du dir Zeit für ein Interview nimmst. Wir kennen uns bereits von meinem Praxissemester und der Freelancing Tätigkeit bei dir, weshalb ich mich dir nicht vorstellen muss. Wie geht es dir?

Markus [00:00:25] Hi, mir geht es gut und ja, sehr gerne.

Lucy [00:00:30] Das freut mich. Wie du bereits aus unseren Vorgesprächen weißt, möchte ich dieses Interview im Rahmen meiner Bachelorarbeit über Besonderheiten der Aufnahmetechnik bei Sprachaufnahmen für Hörbuchproduktionen im Tonstudio verwenden. Ich freue mich, dass du dich dazu bereit erklärt hast, dass ich unser Gespräch aufnehmen und anschließend vereinfacht transkribieren darf. Meine Arbeit inklusive der Transkription kannst du im Laufe der nächsten Monate auf der Webseite von Herrn Curdt abrufen.

Das Ziel des Interviews ist es, mehr über Praxiserfahrungen aus dem Bereich der Hörbuchproduktion zu erfahren. Welche Aufnahmetechnik man für Sprachaufnahmen bei Hörbuchproduktionen im Tonstudio verwenden kann und wie man diese praktisch umsetzen kann, ist dabei die zentrale Fragestellung. In der Arbeit werden Phänomene bezüglich der Aufnahmetechnik erklärt. Daraus soll hervorgehen, warum man auf eine bestimmte Art und Weise bei Sprachaufnahmen im Studio vorgeht und warum man sich für spezifische Aufnahmetechnik entscheidet. Hierbei werden die Themen Schall, Grundlagen der Spracheerzeugung, Mikrofonierung und analoge Echtzeitbearbeitung beziehungsweise Beeinflussung untersucht. Wie man mit der entsprechenden Technik eine gute Sprachaufnahme umsetzen kann und warum, soll in unserem Gespräch deutlich werden. Gerne würde ich anhand der Themen meine Fragen durchgehen. Deine Antworten müssen nicht allumfassend sein. Stattdessen ist mir wichtig, essenzielles und tagtäglich präsent Wissen zu belegen.

Das erste Kapitel meiner Arbeit beschäftigt sich mit Schall und welche grundlegenden Größen wichtig sind, wenn man im Audibereich arbeitet. Daher meine erste Frage: **Welche Grundlagen, physikalische oder akustische, brauchst du in deinem Alltag bei der Sprachaufnahme im Tonstudio? Gibt es ein Thema, das dir ständig präsent ist? Gibt es eine Gleichung, die dir immer präsent ist? Oder gibt es irgendeine Gesetzmäßigkeit in Bezug auf Schall, die du ständig brauchst und die dich, seitdem du etwas darüber gelernt hast, nie mehr verlassen hat?**

Markus [00:03:10] Ich würde nicht sagen, dass man mit Formeln im Kopf im Tonstudio steht. Das macht wahrscheinlich niemand. Aber klar, im Prinzip gibt es wichtige Kenngrößen. Eine der wichtigsten Größen ist sicherlich die Nachhallzeit in Räumen, besonders weil die Sprache für Hörbuchproduktionen raumlos klingen soll. Über den kompletten Frequenzbereich soll ein möglichst trockener Raum im Studio garantiert sein. Der Raum soll so gut wie möglich weggenommen werden, um die Sprachaufnahmen möglichst trocken zu gestalten. Damit einhergehend sind auch Reflexionen ein Thema. Die Eigenschaften von Schall, welche Flächen reflektieren welche Frequenzen. Speziell im hohen Frequenzbereich sorgen zum Beispiel Glasscheiben in Aufnahmerräumen oder auch Displays, auch ein iPad oder ein Tisch für Reflexionen. Alle Flächen, die es um den Sprecher gibt, können zum Problem werden. Und deswegen sind für mich hauptsächlich Nachhallzeit und Reflexion im Raum die Größen, die im Hörbuchbereich wichtig sind. Man berechnet zwar nichts im Alltag, aber man versucht die Voraussetzungen so zu schaffen, sodass die Nachhallzeit und auch Reflexion möglichst gering ausfallen. Zum Beispiel mit Absorbern und Diffusoren.

Lucy [00:05:06] Die Akustik eines Tonstudios mit seiner geringen Nachhallzeit und möglichst wenigen Reflexionen spielt also in Bezug auf Schall eine große Rolle. Im nächsten Kapitel wird die Spracherzeugung oder das akustische Sprachsignal thematisiert. Zum Beispiel gibt es diese Plosive, S- und Zischlaute. **Was gibt es zu beachten bei bestimmten Lauten? Was sind deine Maßnahmen, wie man mit entsprechenden Lauten umgehen kann, sodass sie sich harmonisch in die Sprachaufnahme integrieren?**

Markus [00:06:02] Eine der klassischen Maßnahmen ist die Auswahl des Mikrofons und die Positionierung des Mikrofons.

Weil du die Popplaute angesprochen hast, das ist zu lösen, indem man einen Poppschutz oder Popkiller einsetzt. Das ist die allererste Maßnahme, die immer bei Sprachaufnahmen eingesetzt wird. Bei größeren Mikrofonabständen verringert sich die Intensität des Luftstroms, weshalb Popplaute dann weniger störend erscheinen. Sprachaufnahmen finden allerdings nahe am Mikrofon statt, weshalb Maßnahmen für Popplaute erforderlich sind. Viele Laute hören sich außerdem zu hart an, wenn mit kleinem Abstand direkt in das Mikrofon gesprochen wird. Eine Lösung dafür ist, dass man das Mikrofon seitlicher oder höher positioniert. Im extremen Fall hinter dem Kopf wären allerdings alle S-Laute komplett entschärft, was man nicht wünscht. Infolgedessen macht eine Mikrofonierung auf der Kopfachse, die zu indirekt auf den Mund trifft für Sprachaufnahmen keinen Sinn. Für den subjektiv besten Klang muss man ein bisschen rumprobieren, bis einem die Laute passen.

In Bezug auf die Mikrofonauswahl kann ich sagen, dass unterschiedliche Mikrofone die Sprache und damit einzelne Laute in der Aufnahme verschieden klingen. Zum Beispiel kann ein U 87 von Neumann zu einem brillanten Klang verhelfen, aber in gewisser Weise auch viel Schärfe mit sich bringen. Das tut einer Person, die wie ein scharfes Messer klingt, nicht gut. In dem Fall nehme ich lieber ein OC818 oder sogar so ein dynamisches Mikrofon. Das SM7 beispielsweise wandelt scharfe Laute weniger stark um. Damit kann die Wahl eines geeigneten Mikrofons den Stimmklang ein Stück weit korrigieren. Außerdem kann eine Echtzeitbearbeitung mit analogen Geräten hinter dem Vorverstärker wie ein De-Essern, Equalizer, Kompressoren, ein Hochpassfilter und so weiter zu einer guten Sprachaufnahme beitragen.

Lucy [00:09:53] Wow, das hat gleich mehrere meiner Fragen auf einmal beantwortet. Wenn wir schon bei guten Sprachaufnahmen sind, auch ich habe während meiner Tätigkeit in deinem Tonstudio Unterschiede zwischen Sprechern oder Stimmen bemerkt. **Inwieweit unterscheiden sich geschulte Sprecher von Laien, besonders hinsichtlich der Sprachqualität für dich?**

Markus [00:10:24] Das ist eine spannende Frage, die ich noch nie aktiv beantwortet habe, aber mir selbst auch stelle, beziehungsweise meinen Ohren. An meiner Stimme fällt mir zum Beispiel das Nuscheln auf. Bei geschulten Sprechern kommt das nicht vor, weil sie die Wörter deutlicher aussprechen. Laien hingegen sprechen in der Regel ohne diese sprachtechnisch präzise Fähigkeit und in weniger exaktem Hochdeutsch. Es ist schwierig sich in diesem Bereich fachlich korrekt auszudrücken, aber ich persönlich würde sagen, dass professionelle Sprecher jedes noch so kleine Wort sauber ausdrücken und nichts unbetont lassen. Außerdem ist ein guter Sprecher beim Lesen immer mindestens drei Sätze voraus. Im Vergleich zu durchschnittlichen Leser ist das ein großer Zeitpuffer! Das bedeutet, wenn ein Sprecher vorliest, hat er sich bereits eine passende Betonung für den zu lesenden Text ausgedacht. Aber das ist rein subjektiv. Allerdings sind sich viele einig, dass ein Unterschied zwischen geschulten Sprechern und Laien zu hören ist.

Lucy [00:12:25] Sind dir auch Unterschiede aufgefallen, wenn es um Nebengeräusche geht?

Markus [00:12:30] Ja auf jeden Fall. Klick- und Schmatzgeräusche, Atmen und Bewegungsgeräusche fallen bei beruflichen Sprechern in der Regel gering aus. Wer als Sprecher tätig ist achtet auch auf Kleidung, die keine lauten Reibegeräusche machen. Auch, dass Sprecher keinen Kaffee vor der Aufnahme trinken, ist geläufig. Milchkaffee verklebt den Gaumen und führt zu mehr Klicks in der Aufnahme.

Lucy [00:13:08] Du hast es vorhin schon angesprochen, aber vielleicht noch mal kurz zusammengefasst: **Wie hört sich für dich eine optimal aufgenommene Sprachaufnahme für Hörbuchproduktionen an?**

Markus [00:13:28] Für mich ist das eine Aufnahme, die trocken ist. Bei Hörbuchaufnahmen höre ich keinen Raum. Sie ist verdichtet und trotzdem noch dynamisch genug. Das heißt sie ist nicht zu Tode komprimiert wie Werbung. Warm und brilliant, ähnlich wie Hi-Fi Sound, finde ich klingt ausgewogen. Zwischen den Sätzen und den Sprechpausen höre ich keinen Lüfter oder einen Raum.

Da ist für mich eine angenehme Ruhe. Ich höre keine Klicks, keine Schmatzer. Im Ausnahmefall höre ich reduzierte Mundgeräusche, so dass es natürlich klingt. Das ist für mich eigentlich eine perfekte Sprachaufnahme.

Lucy [00:14:11] Das finde ich auch! Nun kommt ein kleiner Themenwechsel, aber das nächste Phänomen trägt bei fast allen Hörbuchaufnahmen zu einem volleren Klang bei. **Inwiefern setzt du den Nahbesprechungseffekt in der Aufnahme ein? Wann verbessert er die Aufnahme oder verbessert er immer die Aufnahme?**

Markus [00:14:52] Er verbessert nicht immer die Aufnahme. Das ist vor allem davon abhängig, welche Aufnahmetechnik ich verwende und welcher Sprecher spricht. Auch inhaltliche Aspekte spielen eine Rolle für den Naheffekt als gestalterisches Mittel. In der Regel trägt der Effekt zu einer warmen, voluminösen, männlichen Stimme bei, die Hörbuchhörer auch erwarten. Ein guter Sprecher, gute Sprecherin, kann mit dem Naheffekt umgehen. Bei intimen Szenen oder gefühlvollen Inhalten wird der Nahbesprechungseffekt oft bewusst eingesetzt. Die Tiefenanhebung suggeriert mehr Wärme und die Stimme scheint lauter, da leise bei geringerem Mikrofonabstand aufgenommen wird. In der Werbung oder bei Kino Trailern ist das auffällig. Beim durchschnittlichen Hörbuch finde ich einen gesunden Abstand zum Mikrofon passend. Dementsprechend ist der Naheffekt dann nicht mehr so ausgeprägt. Als gestalterisches Mittel wird er berechtigterweise eingesetzt.

Lucy [00:17:21] Folgendes ist demnach vom Inhalt abhängig, aber **was ist deiner Meinung nach ein idealer Abstand zum Mikrofon?**

Markus [00:17:33] Etwa 20 bis 30 cm vom Sprecher bis zum Mikrofon ist für mich optimaler Richtwert für den Abstand. Je nach Sprecherperformance und inhaltlichen Aspekten ist der Abstand wie gesagt individuell.

Lucy [00:17:57] Damit kommen wir auch schon zum Kapitel Mikrofonierung. **Lohnt sich der Einsatz von Mikrofonen mit umschaltbarer Richtcharakteristik?**

Markus [00:18:10] Man braucht nicht unbedingt ein Mikrofon mit umschaltbarer Richtcharakteristik, weil stattdessen auch mehrere Mikrofone mit unterschiedlicher Richtcharakteristik vergleichbar sind. Bei einem Vergleich wird man feststellen, dass sich nicht alle Charakteristiken für Sprachaufnahmen eignen. Beispielsweise würde ich eine Kugel in unserem Studio nicht verwenden. In der Regel werden Nieren für Sprachaufnahmen gewählt, denn dann wird der Raum etwas ausgeblendet. Das trägt zu einer trockenen Aufnahme bei.

Das OC818 ist für mich hinsichtlich der verstellbaren Richtcharakteristiken eine wunderbare Spielerei. Mit dem Plug-In des Mikrofons, in dem man verschiedene Charakteristiken in unterschiedlichen Frequenzbereichen einstellen kann, bieten sich neue Möglichkeiten den Klang der Aufnahme zu beeinflussen. Besonders der Unterschied zwischen super Niere, Niere und breiter Niere ist für Hörbuchaufnahmen interessant.

Je nach Stimme und Raum können bestimmte Richtcharakteristiken in spezifischen Frequenzbändern Vorteile haben. Toll ist außerdem, dass stufenloses Einstellen über den Frequenzbereich möglich ist. Trotzdem würde ich auf keinen Fall sagen, dass man ein Mikrofon mit umschaltbarer Richtcharakteristik braucht. Pauschal empfehle ich für Sprachaufnahmen eine Niere. Alles andere ist nice to have!

Lucy [00:20:40] Es gibt ja auch eine frequenzabhängige Richtcharakteristik. Wann spielt die eine Rolle bei Sprachaufnahmen im Studio?

Markus [00:20:53] Das ist eine gute Frage, immerhin bewegt sich der Sprecher bei Hörbuchaufnahmen nicht sonderlich viel. Wenn der Sprecher allerdings seinen Kopf vom Mikrofon abwendet, kann die frequenzabhängige Richtcharakteristik ausschlaggebend sein. Daher sollte eine breitere Richtcharakteristik für Aufnahmen mit Sprechern eingesetzt werden, die ihren Kopf während der Aufnahme oft drehen. Vereinzelt lesen Sprecher heutzutage noch von ausgedruckten Seiten. Beginnt der Sprecher links oben und endet rechts unten auf einer Doppelseite beim Lesen, sollte unbedingt eine breite Richtcharakteristik verwendet werden. Ansonsten klingt die Sprache immer anders, wenn der Sprecher nicht direkt ins Mikrofon spricht. Um klangliche Veränderungen auf Grund von Kopfbewegungen zu vermeiden, lesen die meisten Sprecher von einem Tablet ab. Außerdem kann mit einem größeren Mikrofonabstand ein kleinerer klanglicher Unterschied bei Kopfdrehungen erzielt werden.

Lucy [00:21:32] Dann kommen wir nun zu etwas technischeren Themen. In der Arbeit schreibe ich über den **Übertragungsfaktor und die Wandlerempfindlichkeit. Wo begegnen dir diese Themen in der alltäglichen Praxis?**

Markus [00:21:57] Aktiv beschäftige ich mich nicht tagtäglich mit diesen Themen. Trotzdem ist es wichtig, dass man die technischen Hintergründe der Aufnahmetechnik kennt. Beispielsweise wenn ich meine analoge Übertragungskette im Tonstudio regelmäßig einmesse, kommt es vor, dass ich solchen Themen begegne. Dabei schicke ich einen Sinuston durch meine Übertragungskette, um zu kontrollieren, ob 0 dB am Vorverstärkerausgang nach der kompletten Kette immer noch die 0 dB sind. Das ist insofern wichtig, dass wenn ein Mikrofon mit der Aussteuerung -6 dB angeschlossen wird, es in der DAW dann auch -6 dB sind. Alle Geräte in der Übertragungskette müssen elektrisch aneinander angepasst sein. Im Regelfall sind die meisten professionellen Studiogeräte das auch, sodass alles pluck and play mäßig funktioniert. Aber einmal alles zur Sicherheit durchmessen lohnt sich. Immerhin nimmt man im Studio fast jeden Tag auf.

Lucy [00:23:19] Eine weitere technische Größe, die ich im Rahmen der Arbeit untersucht habe, ist die **Nennabschlussimpedanz**. **Wann begegnet dir diese Größe?** Der Sprecher meiner letzten Aufnahme hat mich darauf aufmerksam gemacht, dass unterschiedliche Widerstände an Geräteausgängen Auswirkungen auf den Klang der Sprachaufnahme haben. Er hat mir demonstriert, dass er an seinem analogen Gerät verschiedene Abschlussimpedanzen einstellen kann. Klanglich haben die Einstellungen den Unterschied gemacht, dass bei 600 Ohm alles wie gewohnt klang. Bei weniger Widerstand fiel der Klang dünner und subjektiv leiser aus.

Markus [00:25:15] Oh, das ist schon sehr elektronisch! Ich kann dazu sagen, dass man in der Tontechnik immer eine Leistungsanpassung aller Geräte macht. Impedanzen begegnen uns im Bereich der Sprachaufnahme hauptsächlich beim Anschluss vom Mikrofon zu einem Vorverstärker und beim Monitoring über Kopfhörerverstärker und Kopfhörer. Auch hier versucht man eine Leistungsanpassung zu machen. 600 Ohm zu 600 Ohm wäre eine perfekte Leistungsanpassung von eins zu eins, wie zum Beispiel an dem Vorverstärker, den ich habe. An meinem Vorverstärker kann man die Impedanz auch umstellen. Die Größe der Impedanz muss man an das verwendete Mikrofon anpassen. Auch bei diesem Thema ist es nicht so, dass man sich jeden Tag damit beschäftigt, aber als Tontechniker hat man auf jeden Fall mal davon gehört.

Lucy [00:27:43] Dann ist es schön, dass auch ich einmal von einem Toningenieur davon gehört habe. Bisher habe ich mich nur über Bücher mit dem Thema beschäftigt. Nun zu einem alltäglicheren Thema. **Was sind deine Lieblingsmikrofone oder die Mikrofone, die du am meisten einsetzt und warum?**

Markus [00:28:07] Das ist echt eine gute Frage. Ich glaube das variiert über das ganze Leben hinweg. Ich hatte schon viele Mikrofone, die ich gerne eingesetzt habe. In unserem Bereich, wo wir viele verschiedene Sprecher aufnehmen und nicht ständig Mikrofone wechseln wollen oder nicht genau wissen, was am besten passt, ergibt es Sinn einen Allrounder zu besitzen. Das OC818 von Austrian Audio, das eine AKG Kapsel nach altem Vorbild enthält, funktioniert für meinen Studiobetrieb universell und verlässlich. Egal wer vor der Kapsel sitzt, ich habe immer das Gefühl, das spielt gut. Es ist bestimmt nicht das allerbeste Mikrofon für jede Stimme, aber es ist universell und das ist optimal. Das U87 von Neumann, klingt für mich spitz. Für eine männliche Stimme eignet es sich in der Regel gut, allerdings können aufgenommene Frauenstimmen mit dem U87 zu scharf klingen. Ich bin auch ein Fan von der K4 Kleinmembrankapsel von Schoeps. Der lineare Frequenzgang sorgt für ausgewogene Aufnahmen, wenn er zur Sprecherstimme passt. Generell gilt, Mikrofone müssen gut zu dem Sprecher oder der Sprecherin passen. Möglich ist da vieles.

Lucy [00:30:15] Da waren jetzt hauptsächlich Kondensatormikrofone dabei. **Warum benutzt du Kondensatormikrofone?**

Markus [00:30:32] Kondensatormikrofone bilden anders ab als dynamische Mikrofone. Oft klingen sie detailreicher, ihr Frequenzgang ist linearer, die Impulsantwort schneller. Das liegt an der Wandlung und an anderen technischen Gegebenheiten.

Lucy [00:31:02] Apropos Impulsantwort: **Was findest du besser Großmembran- oder Kleinmembranmikrofone? Warum?**

Markus [00:31:19] Ich bin der größte Kleinmembran Fan, keine Frage. Aber die Färbung, die große Mikrofonmembranen in der Stimme erzeugen, funktionieren für Hörbuchaufnahmen gut. Man kann auch mit einer Kleinmembrankapsel Hörbücher aufnehmen. Das funktioniert auch hervorragend und klingt spitze. Es ist eben Geschmackssache. Für die Stimme und das Aufnahmegeschehen Hörbuch ist eine große Membran wegen ihrer warmen Färbung passender. Für alles andere, gerade was Instrumente angeht und Mikrofone an sich bin ich ein Kleinmembranfan.

Lucy [00:33:29] Du hast vorhin schon einige Maßnahmen gegen unerwünschten Schall bei Sprachaufnahmen genannt. Es gibt neben problematischen Lauten auch noch den **Tritt- und Körperschall**. **Wie eliminiert du diesen Störschall während der Sprachaufnahme?**

Markus [00:33:54] Da fängt man bei Tisch und Stuhl im Aufnahmerraum an und geht bis zur Kleidung des Sprechers. Alle möglichen störenden Schallquellen beim Lesen, störende Vorrichtungen, sollten aus dem Aufnahmerraum entfernt werden. Dazu gehören beim ein oder anderen Sprecher oder Sprecherin auch die Schuhe. Tieffrequenter Tritt- und Körperschall wird durch geeignete schwingende Stative und Mikrofonhalterungen reduziert oder vermieden. Auch ein Low-Cut verringert den Anteil von tieffrequenten Störschall.

Lucy [00:34:00] Meine nächste Frage kommt wieder aus dem technischen Bereich. Es geht um den Dynamikbereich und den Störabstand. **In Bezug auf Sprachaufnahmen für Hörbücher, welche Rolle spielen die Größen?**

Markus [00:34:26] Größen wie der SNR geben eine Auskunft über die Übertragungseigenschaften und -qualität von Mikrofonen. Ein Mikrofon, das sehr viel Vorverstärkung braucht, um einen brauchbaren Signalpegel zu erzeugen, rauscht mit großer Wahrscheinlichkeit sehr. Und das wollen wir bei Sprachaufnahmen nicht. Das heißt, es ist wichtig, dass man einen hohen Rauschabstand während der Aufnahme hat. Dann rauscht das Signal nicht. Die Aussteuerung sollte in diesem Zusammenhang ebenfalls betrachtet werden. Ein Sprachsignal sollte ausreichend ausgesteuert werden, ansonsten muss man es digital anheben, was wieder Rauschen verursacht. Das sind Sachen, auf die man achten sollte. Die Aufnahme sollte nicht zu sehr rauschen.

Lucy [00:36:40] Dann noch eine Frage zur **analogen Beeinflussung bei der Sprachaufnahme**. **Wie stehst du generell dazu? Findest du, dass es das braucht, wenn man eine Aufnahme macht? In welchem Maß macht sie die Aufnahme besser? Wie?**

Markus [00:37:04] Ja, das ist eine gute Frage, die ich nur subjektiv beantworten kann. Du kennst die Übertragungskette meines Tonstudios. Es gibt einen Kompressor, eigentlich ein kompletter Channel Strip, den man aber seicht benutzt. Alle analogen Echtzeitbearbeitungen während der Aufnahme sind gewissermaßen risikobehaftet. All das was man in dieser Kette macht, ist auf Tape und nicht rückgängig zu machen. Wird zu viel komprimiert bleibt die Aufnahme zu viel komprimiert. Die Signalspitzen sind dann verloren. Aber ich finde man kann mutig sein. Gerade bei Sprachaufnahmen können die harten Spitzen analog weg komprimiert werden. Mir gefällt es, diesen analogen Weg bis zur Wandlung zu haben, und das Signal so weiterzubearbeiten. Genauso glaube ich, dass man auch vom Vorverstärker direkt in das Interface gehen kann und anschließend nur digital weiterarbeiten kann. Ein Vorteil der analogen Echtzeitbeeinflussung ist, dass wenn man die größten Signalspitzen bereits während der Aufnahme abgefangen hat, schon mehr Ruhe in der Wandlung bringt. Daraus folgt, dass das Sprachsignal dichter gewandelt wird, nach dem Motto: Für das leicht vorkomprimierte Signal können mehr Bits effektiv verwendet werden.

Ich würde trotzdem niemals sagen, es geht nur so oder nur so. Ich bin ein Freund analoger Studiogeräte, besonders von Kompressoren und Limitern. Dabei geht es mir aber nur um die feinen Haarspitzen des Signals.

Lucy [00:40:36] Was meinst du genau damit, dass man nur die Spitzen bei der Aufnahme analog bearbeitet? **Wie sind die Einstellungen am Kompressor für eine geeignete analoge Echtzeitbearbeitung während der Aufnahme?**

Markus [00:41:16] Die Einstellungen können je nach Geschmack, Sprecher und Kompressor variieren. Eine sehr schnelle Ansprechzeit, eigentlich so schnell wie möglich, bis maximal zwei Millisekunden, stelle ich häufig ein. Die Release Zeit ist bei mir lang. Das Kompressionsverhältnis liegt bei unseren Sprachaufnahmen zwischen 1:1 und 3,5:1. Das ist eigentlich so das, was ich oft mache. Für die meisten ist diese analoge Bearbeitung sogar unhörbar. So werden lediglich die Signalspitzen weggenommen. Anschließend kann ein Peak Limiter noch die größten Spitzen kappen.

Lucy [00:43:02] Falls ich nochmal mit einem analogen Kompressor aufnehme, werde ich das auf jeden Fall nachprüfen. Jetzt sind wir auch schon bei meiner letzten Frage angelangt. Da ich weiß welche analogen Geräte bei dir im Studio stehen, wollte ich dich fragen **was dir an den analogen Geräten von Rupert Neve gefällt? Warum hast du dich dafür entschieden?**

Markus [00:43:40] Ach, das ist ja auch Geschmackssache! Vieles ist eigentlich Spielerei. Alle Geräte färben in gewisser Weise. Das finde ich spannend. Bei den Rupert Neve Komponenten aus der 500er Serie finde ich gerade den Texturenregler ganz nett. Er bringt mehr Glanz und Präsenz in die Aufnahme. Was da genau passiert, weiß keiner so richtig. Da müsste man tiefer in der Geräteschaltung forschen. In der Praxis höre ich genau hin und stelle diesen Regler subjektiv ein. An sich ist der Mikrofonvorverstärker Portico 511 aus eben der Reihe ein hochwertiges Gerät, man muss das Signal ordentlich verstärken, bis man ein Rauschen hört.

Ansonsten verfügt er über alle für Sprachaufnahmen notwendigen Einstellungen, von denen ich auch Gebrauch mache: Phantomspeisung, Low-Cut und vieles mehr. Trotzdem sollte man nicht vergessen, dass auch andere analoge Geräte, wie von RME Audio, eine hochwertige Signalbearbeitung ermöglichen. Im Vergleich zum Interface trägt ein guter Vorverstärker zu hochwertigen Aufnahmen meiner Meinung nach mehr bei. Deshalb lohnt es sich in ein passendes Gerät zu investieren, egal von welchem Hersteller!

Lucy [00:45:10] Oh wow, vielleicht werde ich mir dann auch mal einen analogen Vorverstärker zulegen. Danke, dass du deine Erfahrungen mit mir geteilt hast. Damit bin ich mit meinen Fragen am Ende. Ich glaube, dass deine praktischen Fachkenntnisse meine Arbeit ergänzen können. Schön, dass du dir Zeit für das Interview genommen hast. Vielen Dank.

Markus [00:45:53] Kein Problem, bei Fragen, melde dich einfach.

Lucy [00:46:04] Danke, das bringt mich weiter. Dir noch einen schönen Tag. Ciao!

Markus [00:46:10] Tschüss. Nichts zu danken!

2. Hörproben von Hörbuchproduktionen

Produktionen von 2020, 2021 und 2022

Die Hörproben befinden sich auf dem beigefügten USB-Stick. Bei den Dateien handelt es sich um die auf verschiedenen Streaming Plattformen veröffentlichten Hörbuchproduktionen. Alle Dateien liegen deshalb im mp3 Format vor. Alle Produktionen wurden im Namen des Tonstudios Tongoetze und dem Verlag dp DIGITAL PUBLISHERS GmbH realisiert. Im Rahmen dieser Bachelorarbeit dürfen die Aufnahmen verwendet werden.