
BACHELORARBEIT

Herr
Alfred Geier

**Untersuchung der Klangfar-
benmanipulation in Extreme-
Metal-Gesängen mithilfe von
Formantanalysetechniken**

2023

BACHELORARBEIT

Untersuchung der Klangfarbenmanipulation in Extreme-Metal-Gesängen mithilfe von Formantanalysetechniken

Autor/in:
Herr Alfred Geier

Studiengang:
Media and Acoustical Engineering

Seminargruppe:
MG17wA-B

Erstprüfer:
Prof. Dr.-Ing. Michael Hösel

Zweitprüfer:
Rundfunkmechaniker Meister Joachim Kiesler

Bibliografische Angaben

Geier, Alfred:

Untersuchung der Klangfarbenmanipulation in Extrem-Metal-Gesängen mithilfe von Formantanalysetechniken.

Hochschule Mittweida, University of Applied Sciences,
Fakultät Medien, Bachelorarbeit, 2023

Abstract

Extrem-Metal-Gesang, der durch atonale Vokalisation gekennzeichnet ist, spielt eine entscheidende Rolle bei der Vermittlung musikalischer Emotionen durch die Manipulation der Stimmfarbe. Die phonetische Theorie der Formanten liefert eine technische Grundlage für die akustische Analyse dieser Stimmausdrücke. Mithilfe der phonetischen Analysesoftware Praat können Formantfrequenzen im Spektrogramm durch Bereiche konzentrierter Energie im Sonagramm identifiziert werden. Durch die Untersuchung der Formantenverteilung können Erkenntnisse über den stimmlichen Einfluss auf verschiedene Klangfarbeneinstellungen in praxisnahen Screaming-Aufnahmen gewonnen werden. Ziel dieser Forschung ist es, den Zusammenhang zwischen Formantmerkmalen und der emotionalen Wirkung extremer Metal-Gesänge zu untersuchen und so zu unserem Verständnis des stimmlichen Ausdrucks in diesem Genre beizutragen.

Inhaltsverzeichnis

Inhaltsverzeichnis	II
Abkürzungsverzeichnis	IV
Formelverzeichnis.....	V
Abbildungsverzeichnis	VI
Tabellenverzeichnis	VIII
1 Einleitung.....	1
1.1 Einleitung & Problemstellung	1
1.2 Forschungsthese & Zielstellung	2
1.3 Aufbau der Arbeit	2
2 Forschungsstand	3
3 Theoretische Grundlagen	4
3.1 Akustische Grundlagen der stimmlichen Klangerzeugung	4
3.1.1 Schallwellen	4
3.1.2 Periodische Schwingungen	5
3.1.3 Aperiodische Schwingungen	6
3.1.4 Quellen-Filter-Theorie	6
3.1.5 Formanttheorie.....	7
3.2 Mittel der akustischen Klanganalyse	13
3.2.1 Darstellung der Wellenform.....	13
3.2.2 Amplitudenspektrum	14
3.2.3 Sonagramm	14
3.3 Gesangstechniken im Metal	15
3.3.1 Stimmregister.....	16
3.3.2 False Cord Screaming	18
3.3.3 Fry Screaming.....	20
3.3.4 Die Lautheit des Screams	21
3.3.5 Erzeugung der Klangfarbe	21
3.3.6 Die Bezeichnung der Klangfarbe.....	23
4 Methodik	25
4.1 Versuchsaufbau	25
4.1.1 Aufnahmeraum	25
4.1.2 Mikrofonierung	25
4.1.3 Teilnehmende Sänger und deren Musikauswahl	26
4.1.4 Kategorisierung der Gesangsaufnahmen.....	27

4.2	Klanganalyse	28
4.2.1	Formant-Tracking in Praat.....	28
4.2.2	Darstellung der Formanten im Vokaldreieck.....	31
4.2.3	Darstellung der Formanten im Kastendiagramm	32
4.2.4	Darstellung des Amplitudenspektrums	33
5	Ergebnisbericht.....	36
5.1	Klanganalyse	36
5.1.1	Formantanalyse S1	36
5.1.2	Formantanalyse S2.....	41
5.1.3	Formantanalyse S3.....	44
5.1.4	Formantanalyse S4.....	46
5.2	Fazit.....	48
5.3	Optimierungsansätze	51
5.4	Schlusswort.....	52
	Literaturverzeichnis	XIV
	Anlagen.....	A-1
	Eigenständigkeitserklärung	XVII

Abkürzungsverzeichnis

CVT	Complete Vocal Technique
dB	Dezibel
EVT	Estill Voice Training
F0	Grundfrequenz bzw. Grundton
F1 - F4	Erster bis vierter Formant
F1_{av} - F4_{av}	Mittelwert des ersten bis vierten Formanten
F1_{max} - F4_{max}	Maximum des ersten bis vierten Formanten
F1_{min} - F4_{min}	Minimum des ersten bis vierten vFormanten
FC	False Cord
FFT	Fast Fourier Transformation
FS	Fry Scream
Hz	Hertz
IPA	International Phonetic Alphabet
IQA	Interquartilsabstand
LPC	Linear Predictive Coding
Q1 - Q4	erstes Quartil bis viertes Quartil
S1 - S4	Sänger 1 bis Sänger 4

Formelverzeichnis

Formelzeichen	Bezeichnung	Einheit
f	Frequenz	Hz
λ	Wellenlänge	m
c	Ausbreitungsgeschwindigkeit	m / s
T	Periodendauer	ms

Abbildungsverzeichnis

Abbildung 1: Ausbreitung einer Schallwelle	4
Abbildung 2: Parameter einer Schallwelle	5
Abbildung 3: Veranschaulichung der Quellen-Filter Theorie	7
Abbildung 4: Die spektrale Energieverteilung einer Gesangsaufnahme. Die Formanten sind als schwarze waagerechte Bänder erkennbar.	8
Abbildung 5: Die Resonanzräume der Formanten	11
Abbildung 6: Die Einordnung der primären Kardinalvokale nach Daniel Jones	12
Abbildung 7: Darstellung der Wellenform eines Audiosignals	13
Abbildung 8: Darstellung des Amplitudenspektrums eines Audiosignals	14
Abbildung 9: Darstellung einer Gesangsaufnahme im Breitband- (links) und Schmalbandsonogramm (rechts)	15
Abbildung 10: Kategorisierung der Gesangsregister. Modal high bezeichnet die Kopfstimme, Modal low die Bruststimme	16
Abbildung 11: Visualisierung des Kehlkopfes	19
Abbildung 12: Der Vokaltrakt	22
Abbildung 13: Ansicht der Formantspuren in Praat. Die Bereiche auftretender Vokale werden markiert.	30
Abbildung 14: Darstellung von Formantmesspunkten im Vokaldreieck	31
Abbildung 15: Darstellung der Formantmessreihen für verschiedene Klangfarben im Kastendiagramm	32
Abbildung 16: Darstellung der Amplitudenspektren von verschiedenen Klangfarben in Praat	33
Abbildung 17: Darstellung der Amplitudenspektren verschiedener Klangfarben mit der Software R.	35
Abbildung 18: Kastendiagramm der F1- bis F4-Datenreihen von S1 für die Kategorien FC Low (grün), FC Mid (blau) und FC High (rot).	37
Abbildung 19: Darstellung von F1- und F2- Datenreihen der verschiedenen Klangfarben von S1 im Vokaldreieck. Die Mittelwerte sind mit farbigen Kreisen markiert	38
Abbildung 20: Vergleich der Amplitudenspektren der drei Klangfarben von S1.	39
Abbildung 21: Darstellung der drei Klangfarben von S1 im Breitband-Sonogramm.	40
Abbildung 22: Kastendiagramm der F1- bis F4-Datenreihen von S2 für die Kategorien FC Low (grün), FC Mid (blau) und FC High (rot).	42

Abbildung 23: Darstellung von F1- und F2- Datenreihen der verschiedenen Klangfarben von S2 im Vokaldreieck. Die Mittelwerte sind mit farbigen Kreisen markiert.....	43
Abbildung 24: Vergleich der Amplitudenspektren der drei Klangfarben von S2.	44
Abbildung 25: Kastendiagramm der F1- bis F4-Datenreihen von S3 für die Kategorien FC Low (grün) und FC Mid (blau).	44
Abbildung 26: Darstellung von F1- und F2- Datenreihen der verschiedenen Klangfarben von S3 im Vokaldreieck. Die Mittelwerte sind mit farbigen Kreisen markiert.....	45
Abbildung 27: Vergleich der Amplitudenspektren Klangfarben FC Low und FC Mid von S3.	46
Abbildung 28: Darstellung von F1- und F2- Datenreihen der Klangfarbe FC Mid von S4 im Vokaldreieck. Der Mittelwert ist mit einem farbigen Kreis markiert. ..	47
Abbildung 29: Darstellung des Amplitudenspektrums der Klangfarbe FC Mid von S4.	48
Abbildung 30: Darstellung der Mittelwerte verschiedener Klangfarben der Sänger unter Berücksichtigung der Formantenverteilung	49
Abbildung 31: Vergleichende Darstellung der F4-Verteilungen der Sänger	50

Tabellenverzeichnis

Tabelle 1: Die Gemittelten Formantlagen der deutschen Sprache.....	9
Tabelle 2: Auswahl der getesteten Mikrofone	26
Tabelle 3: Einordnung der teilnehmenden Sänger.....	27

1 Einleitung

1.1 Einleitung und Problemstellung

Die menschliche Stimme dominiert fast alle Musikkulturen. Beim Singen fungiert die Stimme als Instrument und gleichzeitig als Vermittler semantischer Bedeutungen. Somit verbindet sie musikalische mit inhaltlichen Emotionen, beispielweise einer Geschichte. Psychologische Theorien legen nahe, dass Menschen die Stimme auf kommunikativer Ebene als sehr bedeutsam empfinden und der Gesang eine wichtige Rolle im Hörverhalten des Musikkonsumenten spielt. Die Emotion der Musik wird kurz gesagt durch den Gesang greifbarer.

Während der Klang der meisten Musikinstrumente auf einem bestimmten Resonanzverhalten beruht, welches eine individuelle Klangfarbe hervorruft, kann bei der Stimme dank der Formbarkeit des Vokaltrakts starken Einfluss auf die Klangformung genommen werden. Durch komplexe Artikulationsmechanismen wird die Bildung verschiedener Sprachen und die damit einhergehende Kommunikation komplexer Sachverhalte möglich. Außerdem wird die Klangfarbe der Singstimme speziell durch die Formung der Vokale beeinflusst, um Emotionen im sprachlichen oder musikalischen Kontext zu verstärken.

Besonders im Extreme-Metal werden gesanglich gern die Grenzen des Möglichen ausgereizt. Grundlegend erzeugen Extreme-Metal-Sänger mit verschiedenen Screaming-Techniken einen brachialen, atonalen Klang, welcher meist energiegeladene Emotionen wie Wut, Aggression, Hass oder aber auch Angst und Trauer betont. Da bei diesen geräuschartigen Klängen in den meisten Fällen keine Tonhöhe entsteht, bedienen sich diese Sänger primär der Vokalformung und weiteren Stimmfaktoren, um die Klangfarbe teils extrem zu manipulieren und damit einen klangästhetischen Eindruck zu erzeugen.

Die Klangeigenschaften der menschlichen Stimme lassen sich physikalisch anhand von Energieverteilungen im Frequenzspektrum erfassen. Die digitale Klanganalyse macht es möglich, dem Phänomen der Klangfarbenmanipulation auf technischer Ebene auf den Grund zu gehen. Dabei handelt es sich speziell um das Verfahren der Formanterkennung, wodurch Aussagen über das Klangbild gegeben werden können.

1.2 Forschungsthese & Zielstellung

Forschungsthese:

Untersuchung der Klangfarbenmanipulation in Extreme-Metal-Gesängen mithilfe von Formantanalysetechniken.

Zielstellung:

Ziel dieser Arbeit ist es, die technischen Möglichkeiten der akustischen Klanganalyse, insbesondere der Formantanalyse, zu erforschen, um verschiedene Formen des Screaming-Gesangs zu untersuchen. Darüber hinaus soll beleuchtet werden, mit welchen Mitteln Metal-Sänger ihre Klangfarbe manipulieren und wie sich dies im physikalischen Bereich der Formantlagen im Klangspektrum manifestiert. Durch die Untersuchung der Formantfrequenzen versucht diese Forschung empirische Belege für den Einfluss verschieden wahrgenommener Stimmklangfarben im Kontext von extremem Metal-Gesang zu liefern.

1.3 Aufbau der Arbeit

Zunächst wird kurz der aktuellen Forschungsstand zum Thema der Formantanalyse von Metal-Gesang geschildert.

Der theoretische Teil ist in drei Abschnitte unterteilt. Im ersten Abschnitt werden technische Aspekte vorgestellt, die zum Verständnis der stimmlichen Klangerzeugung auf physikalischer Ebene beitragen. Anschließend folgt ein Überblick über gängige akustische Darstellungs- und Analysemethoden. Der letzte Abschnitt beleuchtet die Produktion von Gesangstechniken im Extreme Metal aus akustischer und gesangswissenschaftlicher Sicht und bietet eine Einführung in umgangssprachliche Terminologien der Metal-Szene.

Im Methodikteil wird zunächst der Versuchsaufbau der praktischen Klanganalyse dieser Arbeit erläutert. Der zweite Teil der Methodik beschreibt den Ansatz zur analytischen Formantenerkennung sowie möglichen Darstellungsformen der gesammelten Daten.

Im Ergebnisbericht werden die Klanganalysen einzelner Sänger präsentiert. Abschließend wird auf Basis der gewonnenen Erkenntnisse ein Fazit gezogen.

2 Forschungsstand

Sprachforscher verwenden häufig Techniken wie die Spektrogrammanalyse, Wahrnehmungsbewertungen oder Computersimulationen, um Stimmformanten in einem musikalischen Kontext zu untersuchen. Darüber tragen interdisziplinäre Ansätze, die Musikwissenschaft, Akustik und Kognitionspsychologie kombinieren, zu einem umfassenderen Verständnis der Beziehung zwischen Vokalformanten und musikalischem Ausdruck bei. Es gibt zahlreiche Studien, die sich mit der Singstimme und deren Formantstruktur in verschiedenen Genres und Gesangsstilen befassen. Diese Studien untersuchen die Formanteigenschaften und deren Beziehung zu Gesangstechniken oder den Einfluss auf die wahrgenommene Klangfarbe und Ausdruckskraft.

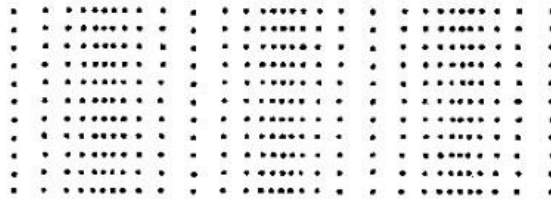
Die Untersuchungen der Formantstruktur im Metal-Gesang ist im Vergleich zu anderen Bereichen der Gesangsforschung relativ begrenzt. Es gibt jedoch einige relevante Studien und Diskussionen zu diesem Thema, welche die Formanten und andere akustische Eigenschaften von Gesangsstilen untersuchen (Smialek E. P., 2012) (Smialek E. D., 2012). Die Formantstruktur von Metal-Gesang bezieht sich auf die Resonanzmuster und Eigenschaften des Vokaltrakts während der Produktion harscher oder extremer Gesangstechniken, die häufig in der Metal-Musik zu finden sind, wie etwa Growls, Screams und gutturale Laute. Andere Studien befassen sich hingegen mit den spektralen Eigenschaften spezifischer Gesangseffekte, die auch im Metal-Gesang ihre Anwendung finden (McGlashan, 2008). Ziel dabei ist häufig der Erkenntnisgewinn über innere Abläufe des Stimmorgans bei diesen Gesangseffekten (Caffier, 2018).

3 Theoretische Grundlagen

3.1 Akustische Grundlagen der stimmlichen Klangerzeugung

3.1.1 Schallwellen

Als Schallwellen werden im Allgemeinen mechanische Schwingungen bezeichnet, die sich in einem elastischen Medium (wie Gasen und Flüssigkeiten) ausbreiten. Dabei entsteht durch die oszillierende Komprimierung und Ausdehnung der Moleküle eine Druckschwankung. Diese Druckänderungen breiten sich in der Luft in Form von longitudinalen Wellen (Längswellen) aus (siehe Abbildung 1).



Longitudinal- oder Dichtewelle

*Abbildung 1: Die Ausbreitung einer Schallwelle
Quelle: Dickreiter, 2013, S.12*

Eine Frequenz einer Welle bestimmt die Anzahl der Schwingungen pro Zeiteinheit gemessen in Hertz. Ein Hertz entspricht einer Schwingung pro Sekunde. Die Amplitude beschreibt die maximale Auslenkung einer Welle aus ihrer Ruhelage.

Der Begriff Wellenlänge bezeichnet bei einer wellenförmigen Ausbreitung den kürzesten Abstand zwischen zwei Punkten der gleichen Phase. Die Wellenlänge λ ergibt sich, indem die Ausbreitungsgeschwindigkeit c im jeweiligen Medium durch die Frequenz f dividiert wird: $\lambda = \frac{c}{f}$.

Eine Schwingung umfasst eine positive und eine negative Halbwelle und dauert eine Periode lang. Die Periodendauer T entspricht dem Kehrwert der Frequenz bzw. dem Verhältnis von Wellenlänge zur Ausbreitungsgeschwindigkeit im jeweiligen Medium:

$$T = \frac{1}{f} = \frac{1}{\frac{c}{\lambda}} = \frac{\lambda}{c}$$

Von einer Welle ist die Rede, wenn sich eine Schwingung im freien Raum ausbreitet, wie beispielweise bei Schallereignissen.

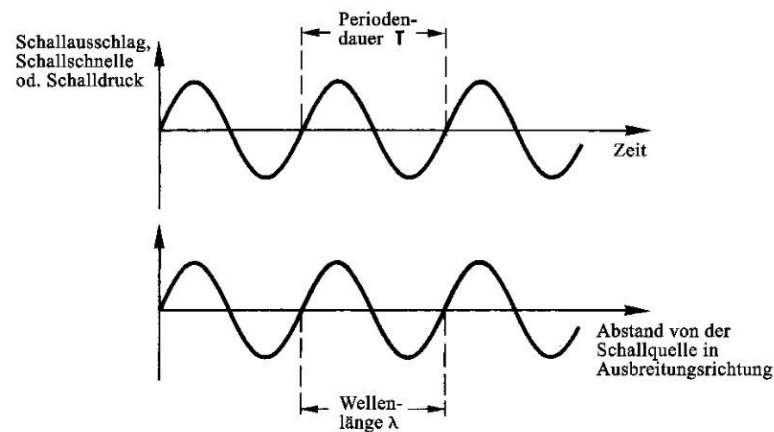


Abbildung 2: Parameter einer Schallwelle
Quelle: Dickreiter, 2013, S.11

3.1.2 Periodische Schwingungen

Die Klangstruktur einer periodischen Schwingung lässt sich in der Akustik durch den Schalldruckpegel des Grundtons und der durch ihn angeregten Obertöne (Partialtöne) beschreiben. Dabei bestimmt der Grundton die wahrgenommene Tonhöhe des Klanges. Die Obertöne sind zusätzlich angeregte Komponenten, deren Frequenzen ganzzahlige Vielfache des Grundtons (harmonische Reihe) bilden. Welche Frequenzen als Obertöne auftreten, hängt von den physikalischen Eigenschaften des jeweiligen Klangereizers ab. Bestimmte Obertonbereiche werden verstärkt, während andere Bereiche gedämpft werden. Die Verteilung der auftretenden Obertöne eines Klanges bestimmt dessen wahrgenommene Klangfarbe. Vokale sind ein Beispiel für periodische Schwingungen der menschlichen Stimme, wobei jeder Vokal eine bestimmte Klangfarbe hervorruft. Harmonische Klänge werden von denen abgegrenzt, die zu den nicht-harmonischen Obertonreihen zählen.

3.1.3 Aperiodische Schwingungen

Bei aperiodischen Schwingungen bilden die Frequenzen der Partialtöne komplizierte, nicht ganzzahlige Verhältnisse zueinander aus. Sie werden tendenziell als Geräusche wahrgenommen und verursachen keine erkennbare Tonhöhe. Daher existiert kein erkennbarer Grundton und die spektrale Frequenzverteilung ähnelt vielmehr einem Rauschen. Aperiodische Schwingungen haben aufgrund dessen keine harmonische Klangstruktur. Geräuschkomponenten der Sprache, welche durch die Zähne, Lippen und die Zunge entstehen, sind Beispiele solcher Schwingungen, etwa stimmlose Laute wie „f“. Diese werden nur durch Strömungsgeräusche der Luft, nicht aber durch das Schwingen der Stimmlippen erzeugt. Allerdings können auch die Stimmlippen aperiodisch schwingen und damit geräuschartige Klänge erzeugt werden.

3.1.4 Quellen-Filter-Theorie

Für Musikinstrumente ist der Klangcharakter weitestgehend durch den Resonanzkörper bestimmt. Bei einer Resonanz wird ein schwingungsfähiges System, beispielweise der Korpus einer Akustikgitarre, mit seiner Eigenfrequenz zum Schwingen angeregt. Je geringer die Dämpfung eines solchen Systems, desto leichter kann es durch Energiezufuhr angeregt werden. Instrumente haben in den meisten Fällen einen unveränderlichen Resonanzkörper, welcher zu einer bestimmten Klangfarbe beiträgt. Die menschliche Stimme kann hingegen bei gleicher Tonhöhe einen sehr unterschiedlichen Klangcharakter ausprägen. Begründet liegt dies in der individuellen Anatomie und besonders der Formbarkeit des Vokaltraktes, wodurch Resonanzen beeinflusst werden können. Ein gleicher Vokal wird bei jedem Menschen unterschiedliche Resonanzen erzeugen und jeder einzelne Vokal ist ebenfalls durch ein bestimmtes Resonanzverhalten des Vokaltrakts charakterisiert. Ein besseres Verständnis der Klangerzeugung der menschlichen Stimme ermöglicht die Betrachtung der Quellen-Filter-Theorie (Humphrey, 2019, p. 84).

Diese stellt die Klangerzeugung als einen zweistufigen Prozess dar, bei dem ein Quellensignal mit der Impulsantwort eines Filters gefaltet wird (siehe Abbildung 3). Die Quelle kann sowohl stimmhaft (periodisch) oder stimmlos (aperiodisch) sein. An erster Stelle der Klangerzeugung der Stimme steht die Energie, die aufgebracht werden muss, um eine Schwingung anzuregen. Diese Energie wird durch die Ausatemluft (Expirationsluft) des Atemsystems gewährleistet, welche die Stimmlippen im Kehlkopf zum Oszillieren anregen.

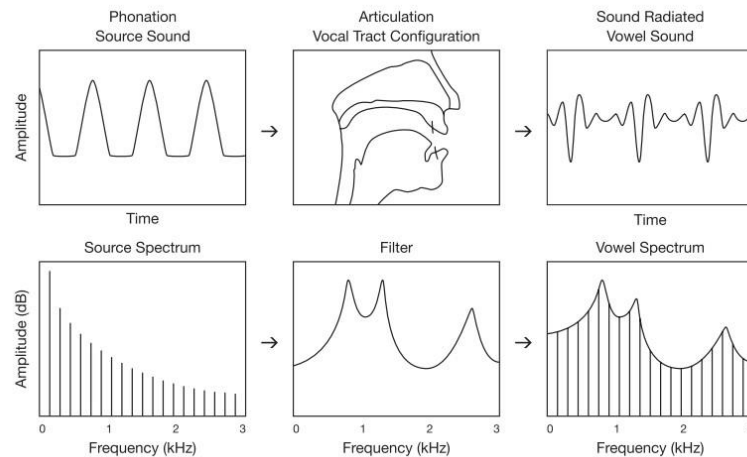


Abbildung 3: Veranschaulichung der Quellen-Filter Theorie

Quelle: Maurer, 2016, p.19

Die Formung des Klanges entsteht durch die anschließenden Resonanzräume des Vokaltrakts, welche durch Mund, Nase, Nebenhöhlen und Rachen gebildet werden (Dickreiter, 2013, S. 66). Dieser wird als unabhängig vom Quellsignal angenommen und bildet einen röhrenförmigen akustischen Resonator, der als Filter wirkt. Die physische Formung unterschiedlicher Laute in den Stimmbändern und der Stimmritze ist als Phonation bekannt, wodurch der Mensch unterschiedliche Phoneme in der Sprache und unterschiedliche Stimmstile im Gesang übermittelt (Humphrey, 2019).

3.1.5 Formanttheorie

Der Charakter einer Stimme wird durch die Grundfrequenz (F_0) und die Artikulationseigenarten bestimmt. Die Grundfrequenz der Stimmlippenschwingung liegt beim Sprechen für Männer zwischen 120 und 160Hz und bei Frauen zwischen 220 und 330Hz (Dickreiter, 2013, S. 68). Oberhalb 3500Hz fällt das Spektrum mit etwa 25dB/Okt stark ab. Dieser Bereich kann durch Betonung zu metallischem Timbre führen. Sehr hohe Frequenzen sind für die Verständlichkeit der Konsonanten wichtig. Dabei reichen stimmhafte Zischlaute bis ca. 8kHz und stimmlose Zischlaute sogar bis ca. 12kHz.

Vokalformanten

Vokale und Umlaute sind Klänge, die durch die obertonreichen Stimmlippenschwingungen im Vokaltrakt Resonanzen erzeugen. Einen Oberton (Partial) oder einen zusammenhängenden Bereich von Obertönen, die bei charakteristischen Eigenfrequenzen durch Resonanzverstärkung im Pegel angehoben werden, bezeichnet man als Formanten. Als Formantbereiche hingegen bezeichnet man Frequenzbereiche,

innerhalb derer unabhängig vom Stimmlippengrundton alle Obertöne durch Vokaltraktresonanzen verstärkt werden (Dickreiter, 2013, S. 66, 67). Wenn also ein Sänger eine Tonleiter auf einem bestimmten Vokal singt, so bleiben trotz der sich ändernden Tonhöhe die Bereiche der stärksten Intensität in ihrer Frequenzlage bestehen (Meyer, 1972). Vokale Emotionen bilden sich in der Klangfarbe der Stimme durch unterschiedliche Hervorhebungen der Formanten bis etwa 3,7kHz aus (Dickreiter, 2013, S. 67). Sie tragen hauptsächlich zur spektralen Hüllkurve der Stimme bei und ändern sich zusammen mit der Länge und Form des Vokaltrakts. Im Vergleich zu den schwingenden Stimmlippen (Quelle) kann der Vokaltrakt (Filter) nur relativ langsame Wechsel aufweisen.

Harmonische, deren Frequenzen nahe an einer Resonanzfrequenz des Vokaltrakts liegen, passieren ungehindert den Vokaltrakt und erzeugen einen Formanten. Harmonische, deren Frequenzen nicht in der Nähe von Resonanzfrequenzen liegen, werden gedämpft. Sie bilden abgeschwächte Bereiche zwischen den Formanten (siehe Abbildung 4).

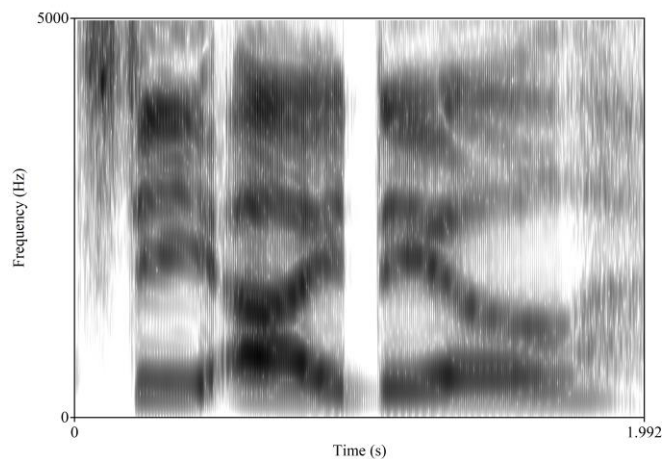


Abbildung 4: Die spektrale Energieverteilung einer Gesangsaufnahme. Die Formanten sind als schwarze waagerechte Bänder erkennbar.

Quelle: Eigene Darstellung

Konsonanten

Bei Konsonanten kommen aufgrund oraler Verengungen Antiresonanzen im Vokaltrakt zustande. Eine Antiresonanz ist das Gegenteil einer Resonanz, sodass die Impedanz, also der Widerstand gegen Vibration bei einer bestimmten Frequenz, bei diesen Frequenzen eher hoch ist (Sundberg J. , The science of the singing voice, 1987, S. 66). Folglich dämpfen oder eliminieren sie Formanten in der Nähe dieser Frequenzen, so-

dass sie abgeschwächt erscheinen oder gänzlich im Spektrogramm fehlen. Hinzu kommt, dass sich der Vokaltrakt für nasale Konsonanten und nasale Vokale in einen nasalen Zweig und einen oralen Zweig teilt (Lewis, 2023). Es entstehen Interferenzen zwischen diesen Zweigen, welche wiederum mehr Antiresonanzen erzeugen. Weiterhin können nasale Konsonanten und nasale Vokale zusätzliche Formanten aufweisen, die durch Resonanz innerhalb des nasalen Zweigs entstehen. Folglich können nasale Vokale aufgrund nasaler Resonanz einen oder mehrere zusätzliche Formanten aufweisen, während ein oder mehrere orale Formanten aufgrund nasaler Antiresonanz geschwächt sein oder fehlen können.

Formantbereiche

Die Formantbereiche sind also Energiekonzentrationen. Diese kann man im Sonagramm als waagerechte Frequenzbänder erkennen und werden mit F1, F2, F3, F4 usw. gekennzeichnet. Während die Vokale hauptsächlich die Lautstärke von Sprache erzeugen, werden über die Konsonanten die Wortdifferenzierungen (Silben) übermittelt.

Vokale äußern sich in ihren akustischen Eigenschaften durch die Lage von F1 und F2, zweitrangig durch die Lage von F3 und geringfügig durch F4 (Lehmann, 2020). Anhand der ersten beiden Formanten lassen sich alle Vokale eines Lautsystems voneinander unterscheiden. Die Vokal-Formantlagen unterscheiden sich von Mensch zu Mensch, besonders zwischen Männern, Frauen und Kindern.

deutscher Vokal	IPA	Formant F1	Formant F2
U	u	320 Hz	800 Hz
O	o	500 Hz	1000 Hz
ä	ɑ	700 Hz	1150 Hz
A	a	1000 Hz	1400 Hz
ö	ø	500 Hz	1500 Hz
ü	y	320 Hz	1650 Hz
ä	ɛ	700 Hz	1800 Hz
E	e	500 Hz	2300 Hz
I	i	320 Hz	3200 Hz

*Tabelle 1: Die gemittelten Formantlagen der deutschen Sprache
Quelle: In Anlehnung an Sengpiel, 2023*

Die gemittelten Formantwerte der deutschen Sprache geben einen ersten Überblick über die wichtigsten Formantbereiche der Vokale (siehe Tabelle 1). Oft werden in den Sprachwissenschaften Vokale und Laute mit einem IPA-Index angegeben. Das IPA (International Phonetic Alphabet) ist ein alphabetisches System der phonetischen Notation, welches versucht, eine standardisierte Darstellung von Sprachlauten in schriftlicher Form zu ermöglichen.

Es ist zu beachten, dass die Klanganalyse im Rahmen dieser Arbeit für englischsprachigen Gesang angefertigt wurde. Dennoch wird zur Einordnung des Klangs in Vokaltendenzen die Notation der deutschen Vokale herangezogen und auf die exakte IPA-Notation verzichtet, um einen vereinfachten und nachvollziehbaren Anhaltspunkt zu bieten.

Sängerformant

Zwischen 2300Hz und 3200Hz bildet sich der sogenannte Sangerformant aus, durch welchen es moglich ist, zwischen individuellen Stimmen zu unterscheiden (Meyer, 1972). Dieser kann geradezu als Qualitatsmerkmal fur Gesangsstimmen gewertet werden, da er der Stimme Tragfahigkeit und Durchsetzungsvermogen gibt (Sundberg J. , The Acoustics of the Singing Voice, 1977). Geubte Sanger manipulieren F3 und F4, indem sie den Kehlkopf senken und das Zungenblatt heben, um diesen Teil des Spektrums zu verstarken und ihn uber einer musikalischen Begleitung horbar zu machen (Lewis, 2023).

Artikulatorische Parameter

Lehmann stellt ein vereinfachtes Modell der Resonanzraume vor, die fur die Erzeugung der Formanten maageblich sind (Lehmann, 2020). Es zeigt sich, dass die Resonanzraume mit steigendem Formant kleiner werden. Der Koronalraum bezeichnet den Raum hinter Oberkiefer und Jochbein (siehe Abbildung 5).

<i>Formantraume</i>				
Formant	Frequenzspektrum (mannlich)			Resonanzraum
F ₀	80	-	200	Stimme
F ₁	220	-	780	Rachen
F ₂	1200	-	2000	Mundraum
F ₃	2200	-	3000	Lippenraum
F ₄	3350	-	4150	Koronalraum

Abbildung 5: Die Resonanzraume der Formanten
Quelle: Lehmann, 2020, Die Sprachlaute I: Vokale

Geschlossene Vokale sind durch einen größeren Rachenraum charakterisiert, wodurch sich F1 tieffrequenter ausprägt. Dies äußert sich durch eine hohe Zungenposition beschreibt dem artikulatorischen Parameter „geschlossen“.

Befindet sich die Zungenspitze vorn, so ist die Rede von dem artikulatorischen Parameter „vorn“. Der Mundraum wird dadurch kleiner. Vordere Vokale lassen sich bevorzugt durch eine größere Differenz zwischen F1 und F2 nachweisen. Zusammenfassend ist für vordere Vokale der Mundraum deutlich kleiner als der Rachenraum, wobei sich für hintere Vokale die beiden Resonanzräume einander annähern.

Die Lippenrundung beeinflusst die Lage von F3, wobei die Rundung der Lippen bewirkt, dass F3 in der Frequenz sinkt. Die Spreizung der Lippen verursacht hingegen einen höheren F3. In den meisten Sprachen sind vordere Vokale gespreizt und hintere Vokale gerundet. Dadurch begleitet ein höherer F3 einen höheren F2 und betont den auditiven Unterschied zwischen vorderen und hinteren Vokalen. Der Verlauf von F3 ist zudem ebenfalls an F1 gekoppelt, da bei offenen Vokalen der Lippenraum kleiner und die Lippenöffnung größer wird, was zu einer Hebung von F3 führt. Vordere gerundete und hintere gespreizte Vokale verursachen hingegen subtilere Unterschiede. F4 ist besonders für Sibilanten wichtig und entsteht, wenn die Korona den Alveolen angenähert wird.

Vokaldreieck

Das Vokaldreieck ist eine Einordnung der F1- und F2-Standardfrequenzwerte aller primären und sekundären Kardinalvokale (Hauptvokale) in einem Koordinatensystem. Daraus ergibt sich die Klassifikation dieser Vokale nach Daniel Jones (1918). Diese ist vielmehr eine idealisierte Darstellung von auditiven Richtwerten, als ein regelmäßiges artikulatorisches System. Das Vokaldreieck zeigt die Extrempunkte der vokalen Artikulation:

Primäre Kardinalvokale

	vorn	hinten	F1
geschlossen	1: i	8: u	250
	2: e	7: o	↓
	3: ε	6: ɔ	
offen		4: a 5: ɑ	750
F2	2200	← 900	

Abbildung 5: Die Einordnung der primären Kardinalvokale nach Daniel Jones
Quelle: Lehmann 2020, Die Sprachlaute I: Vokale

1: Die Zunge ist so weit oben und vorn wie möglich. Dadurch wird ein großer Resonanzraum im Rachen gebildet, was zu einem tiefen F1 führt. Gleichzeitig ist der Mundraum durch die vordere Zungenposition sehr klein, sodass sich F2 hoch ausprägt.

5: Die Zunge ist so weit hinten und unten wie möglich. Dadurch verkleinert sich der Rachenraum und gleichzeitig vergrößert sich der Mundraum. F1 und F2 nähern sich dabei an.

Die Ziffern 1 bis 5 deklarieren Vokale mit gespreizten Lippen, 6 bis 8 sind Vokale mit gerundeten Lippen. Bei den sekundären Kardinalvokalen ist dies genau umgekehrt.

Formantenstimmung

Beim Gesang lassen sich die Formanten ebenso wie bei der Sprache besonders gut für tiefe Töne zeigen. Dies liegt an dem zunehmenden Abstand von Obertönen und der damit verbundenen abnehmenden spektralen Dichte bei höheren Grundtönen. Bereits im höheren Bereich einer Sopranstimme übersteigt die Grundfrequenz für hintere Vokale die Frequenzlage von F1 (Sundberg J. L., 2013, S. 278). Für einen gesungenen Grundton von etwa 700 Hz müssten die Vokale „u“, „ü“ oder „i“, welche die tiefsten F1-Frequenzen aufweisen, demnach unverständlich und schwach wirken. Formanten sind allerdings nicht völlig unabhängig vom Grundton (Sundberg J. , The Acoustics of the Singing Voice, 1977). Wenn der Grundton den Bereich von F1 erreicht oder überschreitet, so steigt F1 zusammen mit dem Grundton an. Diese Anpassung ist auch als Formantenstimmung bekannt und wird dadurch erreicht, dass der Mund beim Singen weiter geöffnet wird. Für vordere geschlossene Vokale, beispielweise „i“ und „e“, sinkt F2 dabei häufig ab. Hintere geschlossene Vokale erfahren hingegen eine gleichzeitige Anhebung von F2. Formanttuning hat seine Grenzen bei einem Grundton von etwa 1000Hz (h2) (Sundberg J. L., 2013, S. 281), sodass für geschlossene Vokale in dieser Tonhöhe der Klangeindruck von F1 völlig verschwindet. Vokale sind in derartig hohen Grundtonbereichen daher nicht mehr gut differenzierbar und die Verständlichkeit verringert sich drastisch.

3.2 Mittel der akustischen Klanganalyse

3.2.1 Darstellung der Wellenform

Die Wellenform stellt die Veränderung der Größe einer Schwingung im Zeitverlauf dar. Für digitalisierte Audiosignale beschreibt diese Größe eine Veränderung der Auslenkung, fachbegrifflich der Elongation der Amplitude, also effektiv die für uns wahrnehmbare Lautstärke. Sie ist durch die Digitalisierung des Audiosignals mathematisch gesehen reduziert auf einen zeitlich sich verändernden reellen Zahlenwert zwischen 0 und 1.

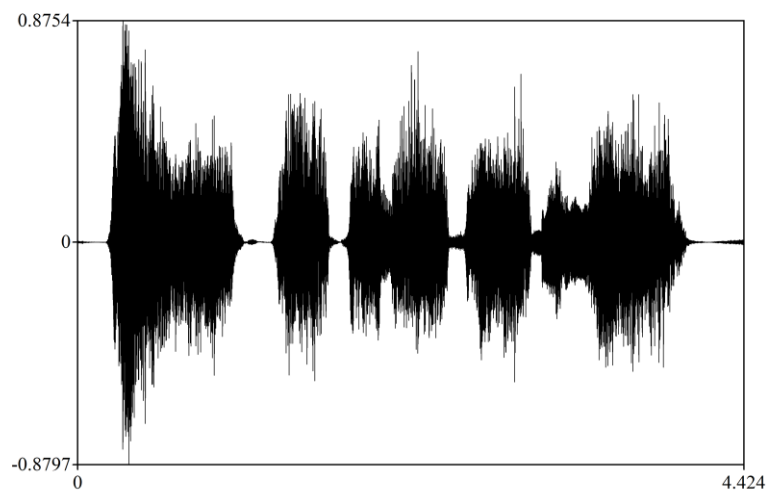


Abbildung 6: Darstellung der Wellenform einer Audiodatei
Quelle: Eigene Darstellung

3.2.2 Amplitudenspektrum

In einem Amplitudenspektrum (auch Frequenzspektrum) können die relativen Amplituden aller Frequenzen eines Signals zusammenfassend dargestellt werden. Dafür wird das Spektrum der Wellenform in kurzen, regelmäßigen Zeitintervallen entlang des Signals analysiert. Die Grundlage, um aus einer Wellenform die Amplituden der Frequenzen zu errechnen ist in der Regel die mathematische Methode der Fourier Transformation. Die Darstellung ist abhängig von dem Zeitabschnitt, der zur Analyse gewählt wird. Der kürzeste Zeitabschnitt muss der Periodendauer der tiefsten Frequenz des Signals entsprechen, um das Spektrum des menschlichen Hörbereichs (20Hz bis 20kHz) vollständig abtasten zu können. Die Periodendauer einer periodischen 20 Hz Schwingung liegt bei 50 ms.

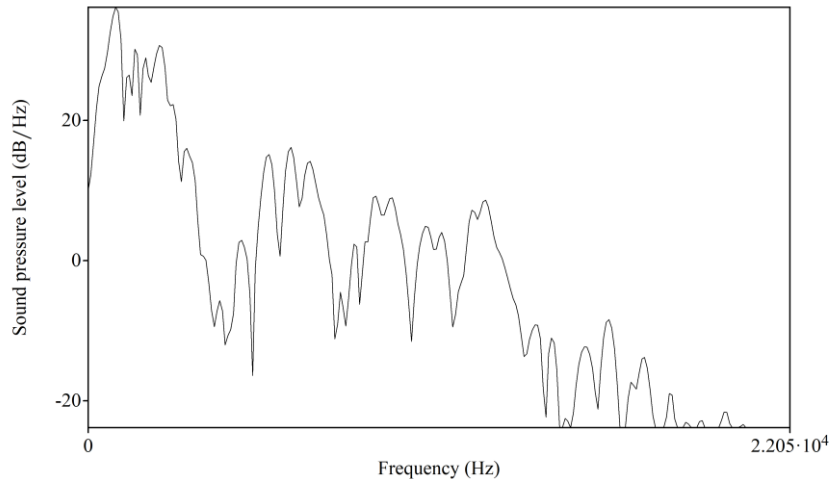


Abbildung 8: Darstellung des Amplitudenspektrums eines Audiosignals
Quelle: Eigene Darstellung

3.2.3 Sonogramm

Als Sonogramm wird ein Schall-Spektrogramm bezeichnet, in welchem die Energieverteilung des Frequenzspektrums über einen Zeitbereich dargestellt wird. Die Stärke der Frequenzkomponenten wird durch den Kontrast der Darstellung sichtbar gemacht. Das Zeitintervall der Abtastung kann über die Filterbandbreite festgelegt werden. Es wird grundlegend zwischen zwei Arten von Sonogrammen unterschieden.

Ein Breitband-Sonogramm zeigt eine hohe zeitliche Auflösung. Das Prinzip der Signalanalyse beruht auf der FFT (fast fourier transformation), die dafür sorgt, dass schnell wechselnde Ereignisse erfasst werden können. Es können in dieser Darstellungsform besonders gut Formantstrukturen der Vokale sichtbar gemacht werden. Ein Formanten äußern sich im Breitband-Sonogramm als waagerechte Frequenzbänder.

Um eine hohe Auflösung im Frequenzbereich zu erzielen, eignet sich das Schmalband-Sonogramm. In diesem sind harmonische Klangstrukturen, also der Verlauf von Grundtönen und derer Teiltonkomponenten besser erkennbar. Diese können in Form von schmalen, waagerechten, meist parallel verlaufenden Bändern identifiziert werden. Der Abstand zwischen den Bändern entspricht dem Verhältnis der Grund- und Teiltöne.

Zugleich eine hohe Zeit- und Frequenzauflösung zu erreichen ist auf Grund des Analysefilters schwierig. Je genauer die Auflösung im Zeitbereich, desto ungenauer ist sie im Frequenzbereich und vice versa (Lewis, 2023).

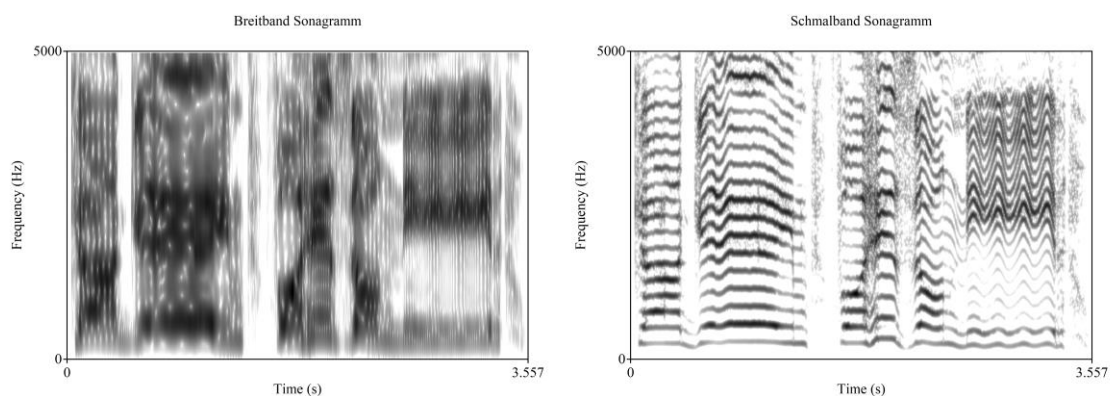


Abbildung 7: Darstellung einer Gesangsaufnahme im Breitband- (links) und Schmalbandsonogramm (rechts)

Quelle: Eigene Darstellung

3.3 Gesangstechniken im Metal

Die Erzeugung der Stimme lässt sich auf eine Reihe von Stufen der Klangerzeugung zurückführen. Complete Vocal Technique (CVT) ist eine international verbreitete Methodik zur Gesangslehre. Die dänische Stimmforscherin Cathrine Salodin beschreibt darin die Klangerzeugung der Stimme in einem vierstufigen Prozess.

Den ersten Schritt bilden drei grundlegende Prinzipien, die besonders auf eine gesunde, schonende und effektive Körperunterstützung abzielen, unter anderem, um mit dem Atem einen konstanten Luftstrom zu erzeugen. Im zweiten Schritt regt dieser die Stimmlippen zum Schwingen an und erzeugt dadurch das Quellsignal der Stimme. Diese Erzeugung der Stimme wird in verschiedene Modi eingeteilt. Diese Kategorisierung stellt eine alternative Herangehensweise zu den bisher in der Gesangswissenschaft definierten Stimmregistern dar. Gesangseffekte können im dritten Schritt dem zugrundeliegenden Modus oder Register hinzugefügt werden. Die letzte Stufe beinhaltet die Beeinflussung der Klangfarbe. CVT ordnet die Wahl des Gesangseffekts aus methodischen Gründen hinter dem Schritt der Klangfarbeneinstellung an. Anatomisch betrachtet finden Gesangseffekte allerdings vor der Manipulation der Klangfarbe statt.

3.3.1 Stimmregister

Unter einem Stimmregister versteht man eine durch dasselbe mechanische Prinzip erzeugte homogene Tonreihe. Alle Töne eines Registers sind folglich gleicher Natur, egal welcher Änderung der Kraft oder Klangfarbe diese unterworfen werden (Roubeau, 2009). Jedes Stimmregister deckt einen bestimmten Tonbereich ab. Diese Bereiche können sich auch überschneiden.

Heutzutage ist man sich in der Gesangswissenschaft allgemein einig, dass Register a) der Tonhöhe und b) der glottalen Stimmquellenkonfiguration – also der Position, Länge und Spannung der Stimmlippen – zugeordnet sind. Jedes Register wird mit einem bestimmten Satz von Stimmlippenschwingungseigenschaften erzeugt. In der Phonetik*, die sich unter anderem der allgemeinen Funktionsweise der menschlichen Stimme widmet, werden meist vier verschiedene Register identifiziert, die klanglich voneinander zu unterscheiden sind (Mesiä, 2015):

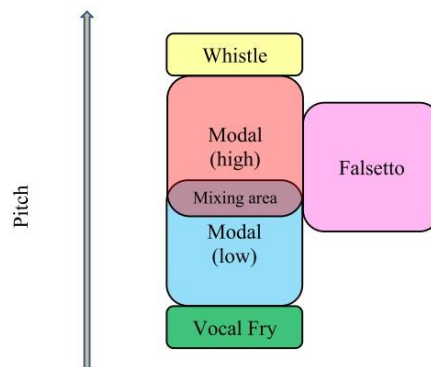


Abbildung 8: Kategorisierung der Gesangsregister. Modal high bezeichnet die Kopfstimme, Modal low die Bruststimme
Quelle: Ribaldini, 2019, p.48

- 1) Vocal Fry (oder „Pulsregister“), das niedrigste Register
- 2) Modal, je nach Tonhöhe in Brust- und Kopfstimme eingeteilt
- 3) Falsett(o), das sich im Tonumfang teilweise mit Modal überschneidet
- 4) Pfeifregister (oder „Flageolet“), die höchste Tonlage.

Vocal Fry

Das Fry-Register wird genauer auch als Kehlkopfmechanismus M0 bezeichnet und zeichnet sich durch sehr kurze, dicke und lockere Stimmlippen aus (Roubeau, 2009, S.

431). Wenn eine Person träge oder ohne Unterstützung durch den Körper spricht, entsteht dadurch ein „knallender“ Klang, bei dem jede Schwingung der Stimmlippen separat zu hören ist. Die Stimmlippen vibrieren aperiodisch bei einer nur leicht geöffneten Stimmritze. Vocal Fry wird im Allgemeinen mit dem tiefsten Tonhöhenbereich der menschlichen Stimme (zwischen 20 und 70 Hz) in Verbindung gebracht. Auf Grund der niedrigen Grundfrequenz hat das Fry-Register eine vergleichsweise geringe Durchsetzungskraft.

Modal

In den Stimmwissenschaften bezeichnet Modal das Register, in dem Sprache normalerweise produziert wird. Es wird in mancher Literatur in das tiefere Sprachregister „Brustregister“ und das höhere „Kopfregister“ eingeteilt. Im Übergangsbereich dieser Register können Töne wahlweise im Brust- oder Kopfregister erzeugt werden. Die Überbrückung zwischen beiden Formen in diesem Bereich wird als Mischen bezeichnet. Das Sprachregister weist eine dicke Kontakttiefe der Stimmlippen auf, was einen vollmundigen, satten Klang erzeugt (Ribaldini, 2019, p. 51).

Die Bruststimme trägt ihren Namen, da sich Frequenzen des tiefen Tonbereichs leichter auf ausgedehnten Oberflächen und großen Körperhöhlen ausbreiten und die dort ansässigen Knochen zu sympathischen Schwingen anregen, einschließlich dem Brustbereich des Sängers. Dieser Effekt kommt besonders bei Männern zum Tragen. Der gleiche logische Ansatz gilt für den oberen Teil des Modalregisters, die „Kopfstimme“. Um den Wechsel zwischen beiden Registern möglichst glatt und kohärent klingen zu lassen, können die beiden Register gemischt werden.

Falsett

Estill Voice Training (McDonald Klimek, 2009, S. 21) definiert „Falsetto“ als Stimmqualität, die dadurch gekennzeichnet ist, dass die Stimmlippen ganz oder teilweise getrennt und in einem steifen Zustand schwingen.

Bei dieser Form der Stimme mit hellem Timbre ist akustisch die Energie der ersten Harmonischen vorherrschend. Bei Männern ähnelt es der Jungenstimme. Der Unterschied zum Modal ist somit größer als bei Frauen, da bei ihnen die pubertären Stimmveränderungen nicht so drastisch ausfallen. Im unteren Bereich klingt Falsett gehaucht. Wenn es höher ansteigt wird der Hauch reduziert und es ist schwieriger von der Kopfstimme zu unterscheiden.

Pfeifregister

Das Pfeifregister ist noch recht unerforscht und die Literatur darüber noch nicht umfangreich. Es ermöglicht die Erzeugung der höchsten Töne in der menschlichen Stimme, wird aber weniger in Sprachphonation oder Gesang angewendet und nur eine Minderheit von Sängern weiß, wie sie darauf zugreifen können (Roubeau, 2009, S. 429). Es ist durch eine schmale Öffnung der Stimmlippen gekennzeichnet, welche sehr dünn und gestreckt sind um einen Pfeifton zu ermöglichen.

3.3.2 False Cord Screaming

Gesangstechniken, die häufig ihre Anwendung im Metal-Bereich finden, bestehen zu meist nicht nur aus einem Effekt, sondern aus Effektkombinationen (Linke, 2023). Um im späteren Praxisverlauf eine sinnvolle Einteilung der Gesangsaufnahmen treffen zu können, ist eine kurze Einführung in die umgangssprachliche Terminologie der verschiedenen Gesangstechniken anzuführen. Dabei stellen sich zwei Hauptbezeichnungen heraus, welche sich in der Metal-Szene etabliert haben. Zu bemerken ist, dass teilweise physiologisch oder technisch widersprüchliche Begrifflichkeiten verwendet werden. Außerdem sind die meisten Begriffe nicht klar definiert. Die Meinungen von Fachleuten gehen teilweise stark auseinander.

False Cord Scream ist eine Bezeichnung, die häufig verwendet wird, um jeden Schrei zu beschreiben, der hauptsächlich auf den Vibrationen der Taschenfalten bzw. den "False Cords" basiert (Abbildung 9). Diese Technik ermöglicht es dem Metal-Sänger, eine große spektrale Streuung der Energie zu erzeugen, wie man dies in Sonagrammen erkennen kann (Smialek E. D., 2012, S. 88). Physiologisch bestehen einige Ähnlichkeiten zum Untertongesang, der in einigen ethnischen Stimmtechniken verwendet wird, wie dem Kehlgang der Tuwiner oder dem Gesang tibetisch-buddhistischer Mönche (Mesiä, 2015, S. 388).

Grunt

Als Grundlage des *False Cord Screams* kann der *Grunt* identifiziert werden, welcher die charakteristischen Stimmlaute vieler extremen Metal-Genres beschreibt. Dabei vibriert der gesamte Kehlkopf bei offen positionierten Stimmlippen (siehe Abbildung 9, links), was sehr tieffrequente und kraftvolle, aperiodische Schwingungen verursacht (Salodin, *Complete Vocal Technique*, 2008, p. 196). Da die Stimmlippen nur geringfügig an der Klangerzeugung beteiligt sind, entsteht bei diesem Effekt kein erkennbarer

Grundton. Infolgedessen ist *Grunt* bei Männern und Frauen oft schwierig voneinander zu unterscheiden. Dieses Phänomen hat vielen Sängerinnen den Weg hin zu einer Sängerkarriere in Extreme Metal Bands geebnet.

Grunt kann im Kontext des *False Cord Screams* allein stehend oder in Kombination mit anderen Effekten angewendet werden. Häufig kombinierte Effekte sind dabei bspw. *Distortion*, *Rattle* und *Creaking*.

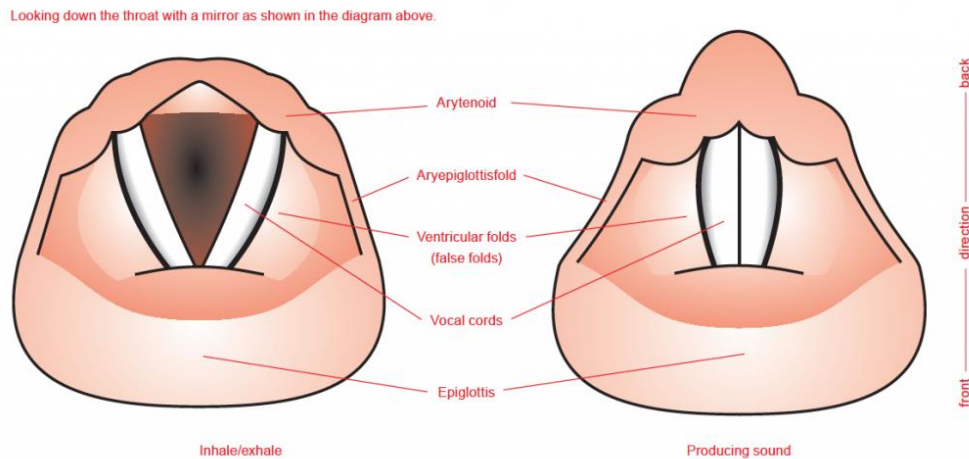


Abbildung 9: Visualisierung des Kehlkopfes

Quelle: Salodin, CVT Research, 2023, *Larynx and the Vocal Folds*

Distortion

Verzerrungseffekte verschiedener Natur werden üblicherweise verwendet, um der Stimme mehr Ausdruckskraft zu verleihen. Der Begriff Verzerrung umfasst in verschiedener Literatur eine Vielzahl von Effekten, die sich sehr voneinander unterscheiden können. CVT meint damit konkret die zusätzliche Vibration der Taschenfalten (Salodin, *Complete Vocal Technique*, 2008, p. 179). Eine solche Schwingung kann aperiodisch oder periodisch sein. Im Metal-Gesang äußert sich die Verzerrung überwiegend aperiodisch in Form von Rauschen.

Growl

Growl ist eine etwas weichere Art der Verzerrung, wie sie beispielweise Louis Armstrong verwendet hat. Dieser Effekt ist oft in Gospel oder Jazz hörbar. *Growl* wird bei normaler Phonation der Stimmlippenebene und mit einem Turbulenzeffekt erzeugt, der durch eine deutliche Neigung der Epiglottis in Richtung der Stellknorpel (arytenoids)

entsteht (siehe Abbildung 9) (Salodin, *Complete Vocal Technique*, 2008, p. 192). Dadurch äußert sich ein „rollender“ Klang. Beim Growl hat der Sound einen deutlich hörbaren Grundton, beim *Grunt* dagegen ist der Grundton oft nicht eindeutig erkennbar.

Rattle

Rattle ist ein Gesangseffekt, der höher im Vokaltrakt erzeugt wird als bspw. Distortion. Der Effekt entsteht durch die Schwingungen der Stellknorpel, teilweise auch des hinteren Teils der Zunge oder des Gaumensegels (Salodin, *Complete Vocal Technique*, 2008, p. 191). In isolierter Form kann Rattle weniger ausdrucksvoll wahrgenommen werden als andere Verzerrungseffekte. In Kombination kann Rattle die Ausdruckskraft anderer Effekte bekräftigen.

3.3.3 Fry Screaming

Der *Fry Scream* ist eine Bezeichnung, die häufig verwendet wird, um jeden Schrei zu beschreiben, dessen Geräuschanteil hauptsächlich auf chaotischer bzw. komplexer Vibration der Stimmlippen basiert. Fry Screams werden gewöhnlicherweise aus einem zugrundeliegenden Modus oder Register und dem Effekt *Creaking* (oft „Fry“ genannt) erzeugt (Salodin, *Complete Vocal Technique*, 2008, p. 187). Nicht zu verwechseln ist der Effekt „Fry“ mit dem gleichnamigen Register, bei welchem ein tiefer Fundamentaltone durch schlaffe Stimmbänder entsteht. Für den Effekt *Creaking* werden die schlaffen Stimmbänder, abhängig vom verwendeten Register, gestrafft, wodurch komplexere aperiodische oder subharmonische Schwingungen erzeugt werden. Weitere Effekte, beispielweise *Distortion* oder *Rattle*, können hinzugezogen werden, um die Lautstärke des Screams zu erhöhen. Als Stimmregister kann beispielweise eine hohe Bruststimme (Belting) gewählt werden, um einen besonders lauten Scream zu erzeugen. In Mischstimme bildet der Fry Scream eine mittlere, im Falsett meist eine niedrige Lautstärke aus (Linke, 2023).

3.3.4 Die Lautheit des Screams

Die Lautstärke eines Screams hängt besonders von der Wahl des Stimmregisters und der Effekte ab. Mögliche Begriffe, um die wahrgenommene Lautheit eines Screams zu beschreiben sind bspw. *quiet*, *medium* und *loud*, verbreitet sind auch die Bezeichnungen „Whisper Screams“ bzw. auf der anderen Seite „Yell Screams“ oder „Shouts“ für die dynamischen Extreme. Eine Technik, die sich besonders abhebt, ist der *hyper-*

compressed Fry Scream. Wie es der Name andeutet, wirkt der Scream energetisch stark komprimiert, wodurch bei vergleichsweise geringer Lautstärke ein druckvoller Klang erzeugt wird. Dieser dynamische Effekt wird durch die Anwendung des Valsalva-Manövers erzeugt, bei dem ein hoher Druck unterhalb der Stimmlippen und Taschenfalten hergestellt wird (Linke, 2023).

Dabei wird versucht, gegen eine geschlossene Stimmritze kräftig auszuatmen, wobei die Atem- und Bauchmuskulatur stark angespannt wird. Dieser Effekt ist auch besser nachvollziehbar, wenn man sich bspw. vorstellt, ein Gewicht zu heben.

3.3.5 Erzeugung der Klangfarbe

Die Klangfarbe der Stimme hängt von der individuellen Physiologie des Vokaltraktes eines Menschen ab. Abgesehen davon kann die Form des Vokaltrakts in verschiedenen Wegen verändert werden, um die Klangfarbe zu beeinflussen. Wobei sich die Formung der Vokale hauptsächlich auf die Position von Zunge und Lippen stützt, können zur Klangfarbenmanipulation außerdem der Kehlkopftrichter, die vertikale Kehlkopfposition, der Gaumen als auch der Nasengang verändert werden.

Einem guten Sänger wird zumeist eine gleichmäßige bzw. kontrollierte Klangfarbe, als auch eine Lautstärkeunabhängigkeit von der Tonhöhe zugeschrieben. Sadolin gibt eine simple Übersicht über die Erzeugung der Klangfarben:

Kehlkopftrichter

Über den Stimmbändern befinden sich zwei viereckige Membranen. Zusammen mit der davor befindlichen Epiglottis und den Stellknorpeln dahinter formen diese einen Trichter, welcher auch der Kehlkopftrichter genannt wird. Bei der Anwendung des Twangs* wird dieser Trichter schmaler, indem die Stellknorpel näher an den unteren Teil der Epiglottis gebracht werden. Je mehr der Tunnel „gequetscht“ wird, desto spitzer und durchdringender wird der Klang (Megafoneffekt). Schon mit einem moderaten Twang kann die Lautstärke um 10 bis 15 dB angehoben werden. Problematisch kann Twang bei dem Vokal „o“ werden, da der hintere Teil der Zunge abgesenkt wird und dadurch den epiglottischen Trichter beeinflusst. Diese Vokale tendieren durch Twang mehr zu einem „e“ („oe“) (Salodin, Complete Vocal Technique, 2008, p. 159).

Kehlkopfhöhe

Die Höhe des Kehlkopfes (Abbildung 10) ist stark abhängig von der gesungenen Tonhöhe. Bei tiefen Tönen ist der Kehlkopf eher tiefgestellt, bei hohen Tönen wird auch der

Kehlkopf angehoben. Allerdings trägt die Position des Kehlkopfes auch in ähnlicher Form zur Klangfarbe bei. Innerhalb des Bereichs eines bestimmten Tons kann die Kehlkopfposition angepasst werden, um Einfluss auf die Klangfarbe zu nehmen. So klingt bei gleicher Kehlkopfhöhe ein etwas tiefer Ton heller und ein etwas höherer Ton dunkler.

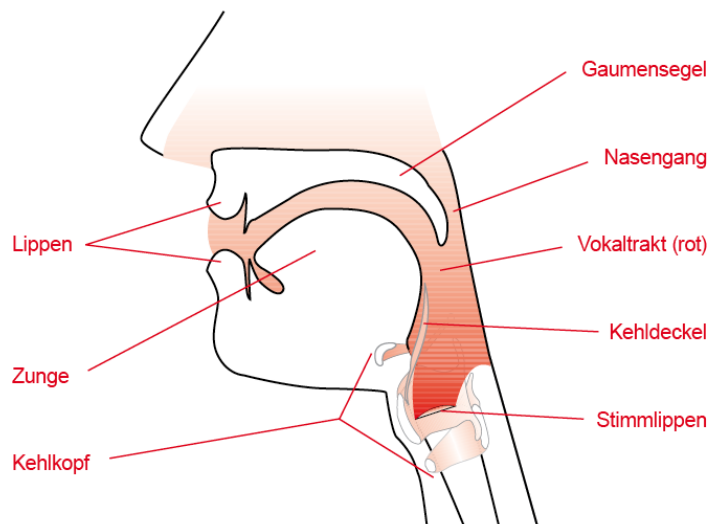


Abbildung 10: Der Vokaltrakt
Quelle: Salodin, CVT Research, 2023, *Description and sound of Sound Colour*

Zungenstellung

Indem die Zunge komprimiert wird, also flach in der Mundhöhle liegt, so kann eine dunkle Klangfarbe erreicht werden. Grund dafür ist, dass dadurch der Vokaltrakt vergrößert wird. Dies ist bei den Vokalen „a“ und „o“ der Fall. Wird hingegen die Mundhöhle schmaler, indem die Zunge an die Backenzähne gedrückt wird und sich zusätzlich Richtung Gaumen biegt, so kann ein heller Klang erzeugt werden. Der Klang wird mehr nach den Vokalen „e“ und „i“ klingen (Salodin, *Complete Vocal Technique*, 2008, p. 166).

Formung des Mundes

Sind die Mundwinkel entspannt (wie bei einem „o“) so wird ein dunkler Klang erreicht, da sich der Vokaltrakt vergrößert. Dehnen sich allerdings die Mundwinkel aus, wie dies beim Lächeln der Fall ist, so entsteht eine hellere Klangfarbe.

Formung des Gaumens

Auch der Gaumen kann in seiner Form variiert werden. Wird der Gaumen angehoben, so bewirkt dies eine Vergrößerung des Vokaltraktes. Oftmals wird dadurch der Kehlkopf automatisch mit abgesenkt, was den Effekt eines dunklen Klanges verstärkt. Die reguläre Stellung des entspannten Gaumens bewirkt im Gegensatz dazu eine hellere Klangfarbe und ist meist mit der Anhebung des Kehlkopfes gekoppelt.

Öffnung des Nasenganges

Ist der Nasengang geöffnet und die Zunge blockiert nicht den Mund, so kommt der Klang sowohl aus Nase als auch aus dem Mund und erzeugt einen nasalen Klang. Wird zusätzlich der Mund durch die Zunge blockiert, so kommt es zu einem rein nasalen Klang, wie bei den Lauten „m“ oder „n“. Ist der Nasengang hingegen vollkommen verschlossen, so erzeugt dies einen kräftigen und dunklen Klang.

3.3.6 Die Bezeichnung der Klangfarbe

Um die beiden vorgestellten Screaming-Techniken in ihrem Klang genauer zu spezifizieren, haben sich in der Szene die Begriffe *Low*, *Mid* und *High Screams* etabliert. Damit ist keinesfalls die Höhe des verwendeten Grundtons gemeint, da diese Techniken zumeist ohne oder mit gering anteiligen Grundtonkomponenten erzeugt werden. An dieser Stelle beschreiben *Low*, *Mid* und *High*, wie dunkel bis hell sich die Klangfarben des erzeugten Screams ausprägen. Dies wird besonders durch die Mittel der gesanglichen Klangfarbenmodifikation erreicht, die bereits vorgestellt wurden und das Ziel haben, die Größe des Vokaltrakts zu verändern. Den größten Einfluss auf die Klangfarbe nimmt vermutlich die vertikale Kehlkopfposition in Verbindung mit der Lippenformung, welche zur Artikulation verschiedener Vokale verwendet wird. Dieser Punkt wirft die wohl wichtigste Technik auf, die Metal-Sängern zur Verfügung steht: Die Manipulation von Vokalformanten. Diese Sänger opfern häufig die Verständlichkeit textlicher Inhalte um eine gewisse Übertreibung der Klangfarbe zu erzeugen, wie beispielweise mit besonders dunklen Screams/tiefen Vokalformanten im Death-Metal-Bereich (Smialek E. D., 2012, S. 89).

Zusammenfassend ist die Klangfarbe von Screaming-Techniken nicht zwingend von den angewandten Effekten selbst, sondern stärker von deren artikulatorischen Variationen abhängig.

4 Methodik

4.1 Versuchsaufbau

4.1.1 Aufnahmezimmer

Die Gesangsaufnahmen entstanden in den Einrichtungen der Firma Musikelektronik Geithain. Als Aufnahmezimmer diente ein 20 m³ großer, akustisch optimierter Raum, der ursprünglich für den Test von Lautsprechersystemen konzipiert war. Der Zweck dieses kleinen Raums bestand darin, die akustischen Bedingungen zu simulieren, die normalerweise in einem Übertragungswagen des Rundfunks mit begrenztem Platzangebot anzutreffen sind. Dieser Raum dient als valide Aufnahmeumgebung, in der Klangverzerrungen durch unerwünschte Raummoden vermieden werden können.

4.1.2 Mikrofonierung

Um ein Höchstmaß an Aufnahmequalität zu gewährleisten, wurden verschiedene Mikrofone getestet (siehe Tabelle 2). Um deren Klangeigenschaften zu vergleichen, wurden diese Mikrofone mit genügend Abstand nebeneinander positioniert, sodass der Sänger nacheinander mit jedem Mikrofon eine identische Passage einsingen konnte. Die Klangqualität wurde im Rahmen des Auswahlprozesses durch einen A-B-Vergleich bewertet.

Signifikante Unterschiede konnten festgestellt werden, wenn jede Aufnahme im gleichen Verhältnis einer starken dynamischen Komprimierung unterzogen wurden, wie es bei der Produktion von Metal-Musik häufig der Anwendungsfall ist. Aufgrund des begrenzten Umfangs dieser Arbeit kann eine detaillierte Darstellung des Vergleichs nicht bereitgestellt werden. Zusammenfassend lässt sich sagen, dass die auffälligsten Unterschiede im Frequenzbereich der hohen Mitten zwischen 1kHz bis 4kHz zu verzeichnen waren. Dieser Bereich ist entscheidend für die Projektion der Stimme, besonders der Durchsetzungskraft und Tragfähigkeit (Sängerformant). Letztendlich fiel die Entscheidung auf das Neumann U87 Mikrofon.

Mikrofon	Physikalisches Prinzip
Lewitt LCT 140	Großmembrankondensatormikrofon
AKG P820	Röhrenmikrofon
Shure SM 7B	Tauchspulenmikrofon
Soyuz 017 FET	Großmembrankondensatormikrofon
Neumann U87	Großmembrankondensatormikrofon
Neumann M520	Kugelmikrofon

*Tabelle 2: Auswahl der getesteten Mikrofone,
Quelle: Eigene Darstellung*

4.1.3 Teilnehmende Sänger und deren Musikauswahl

Zur Durchführung der Klanganalyse wurden Gesangsaufnahmen von vier verschiedenen Sängern im Alter zwischen 23 und 31 Jahren erstellt. Ihre Erfahrungen im Metal-Gesang reichte von 1 bis 11 Jahren. Neben ihrem Erfahrungsniveau variierte auch ihre musikalische Aktivität erheblich. Der am wenigsten aktive Sänger gab an, etwa ein- bis zweimal im Monat an Bandproben beteiligt zu sein, wohingegen der engagierteste Sänger täglich übte, an mehreren musikalischen Projekten beteiligt war und auch Gesangsunterricht erteilte.

Die Sänger wurden angewiesen, eine Auswahl von 4 bis 6 Titeln zu treffen, die ein möglichst diverses Spektrum an Gesangsstilen abdeckt. Drei der vier Sänger wählten ihre Titel mit einem starken Fokus auf genrespezifische Entscheidungen. Sänger 1 zeigte jedoch ein stilistisch vielfältiges Repertoire. Darüber hinaus zeigt Sänger 1 zusammen mit einem anderen Sänger ein breites Spektrum an Gesangstechniken, was wahrscheinlich auf ihre umfangreiche Erfahrung zurückzuführen war. Sie waren in der Lage, verschiedene False-Cord-Screaming-Techniken in den zuvor vorgestellten dunklen bis hohen Stimmfarben sowie verschiedene Fry-Scream-Techniken liefern. Im Gegensatz dazu konzentrierten sich die anderen Sänger hauptsächlich auf spezifische Gesangseffekte.

Die Sänger wurden nach ihrer Erfahrung, Aktivität und der Wahl des Subgenres eingeordnet (siehe Tabelle 3). Eine Einordnung der Subgenres kann sehr mehrdeutig ausfallen und dient nur der Orientierung.

	Gesangserfahrung	Aktivität	Subgenre
Sänger 1 (S1)	11 Jahre	Täglich	Heavy Metal, Progressive Metal, Melodic Metal, Death Metal
Sänger 2 (S2)	10 Jahre	Wöchentlich	Metalcore, Progressive Metal
Sänger 3 (S3)	7 Jahre	Monatlich	Death Metal, Doom Metal
Sänger 4 (S4)	1 Jahr	Mehrmals die Woche	Stoner Rock, Progressive Metal

*Tabelle 3: Einordnung der teilnehmenden Sänger
Quelle: Eigene Darstellung*

Eine ausführliche Übersicht der Songauswahl jeden Sängers ist im Anhang zu finden (siehe Anlage 4).

4.1.4 Kategorisierung der Gesangsaufnahmen

Folgende Kategorien wurden zur Klanganalyse der Tonaufnahmen erstellt:

Kategorie FC: False Cord Scream

In dieser Kategorie kommt die gleichnamige Technik zum Einsatz. Diese konnte von allen Sängern erzeugt werden. Folgende Unterkategorien in Bezug auf die Klangfarbe wurden gebildet:

False Cord Low (FC Low), False Cord Mid (FC Mid) und False Cord High (FC High).

Nicht jede Unterkategorie wurde von allen Sängern abgedeckt. Dies kann auf die Fähigkeiten des Sängers, oder aber auf dessen spezifisches Genre zurückzuführen sein. Im Death Metal finden High Screams beispielweise kaum eine Anwendung.

S1 und S2 konnten False Cord Screams in allen Unterkategorien erzeugen und standen daher im Mittelpunkt der Klanganalyse. Somit konnte bei beiden Sängern die Erzeugung einer dunklen bis hellen Klangfarbe gegenübergestellt werden. S3 bot nur die Kategorien *FC Low* und *FC Mid* mit einem generell dunkleren Klangspektrum. S4 hingegen deckte ausschließlich die Kategorie *FC Mid* ab.

Kategorie FS: Fry Scream

Zwei Sänger lieferten Aufnahmen in der Kategorie *Fry Scream*. S2 führte einen *loud Fry* vor, während S3 einen *hyper-compressed Fry* darbot. Aufgrund des begrenzten Umfangs der Aufnahmen in dieser Kategorie lag der Fokus der Klanganalyse ausschließlich auf der ersten Kategorie.

4.2 Klanganalyse

Wie bereits erläutert sind Amplitudenspektren (Spectral Slices) nützliche Hilfsmittel zum Vergleich lokaler Spektralereignisse oder zum Messen von Spektraleigenschaften wie Formantfrequenzen, Formantpegeln oder Formantbandbreiten. Sonagramme bieten ebenfalls eine Möglichkeit der Darstellung und Analyse von Formantbereichen.

Praat ist eine phonetische Analysesoftware, die es möglich macht, Sprach oder Gesangsaufnahmen anhand der Erstellung von Sonagrammen und deren Auswertung in ihren physikalischen Bestandteilen zu erfassen. Die Hauptfunktion auf der diese Arbeit aufbaut ist das Verfolgen von Formantbereichen im Sonagramm und deren Analyse.

Um die Klangfarbe jeden Sängers und deren Gesangstechniken erfassen zu können, werden die Analysemethoden der spektralen Analyse und der Formantenerkennung kombiniert. Mit den gewonnenen Erkenntnissen wird ein Vergleich angestellt mit dem Ziel, Gemeinsamkeiten und Unterschiede gleicher Modi (*FC Low*, *FC Mid* und *FC High*) bei verschiedenen Sängern zu untersuchen.

4.2.1 Formant-Tracking in Praat

Für die Formantenverfolgung (formant tracking) werden die Formanten in regelmäßigen Zeitintervallen entlang des Signalpegels aufgezeichnet, wodurch eine kontinuierliche Folge von Messpunkten durch jeden Formanten erzeugt wird. Für eine zuverlässige Erfassung des Pegels kommt das LPC (linear predictive coding) Verfahren zum Einsatz, welches nicht von der Stimmlage beeinflusst wird (wie dies bei FFT der Fall ist) und daher erfolgreicher beim Auffinden der Formanten in einer höheren

Stimmlage ist. Ihre Anwendung findet die Formantenverfolgung bei Vokalen, da deren Resonanzbereiche gut erkennbar sind und korrekt identifiziert werden können.

Das Verfahren beginnt zunächst mit der Erstellung eines Spektrogramms, fährt dann damit fort, die Formanten gemäß der Parametereinstellungen zu finden und sie dann auf dem Spektrogramm darzustellen. Um die Formant-Spur zu optimieren, müssen einige Parameter fein eingestellt werden. Dazu muss die Anzahl der zu suchenden Formanten definiert werden und die erwartete maximale Frequenz des höchsten dieser Formanten. Die Werte dieser Parameter sind stark personenabhängig und spiegeln die individuelle Vokaltraktlänge wider.

Wenn die maximale Frequenz zu niedrig eingestellt ist, werden ein oder mehrere Formanten ausgeschlossen und nicht verfolgt. Wenn sie zu hoch eingestellt ist, sind während der Suche mehr Formanten als erforderlich vorhanden, was zu einer möglichen Mehrdeutigkeit bei der Identifizierung von Formanten führt. Abbildung 11 zeigt ein Sonagramm und dessen Tracking für F1 bis F4. Mit Hilfe des darunterliegenden Textfeldes werden die Bereiche der Aufnahme ausgewählt, in denen das Formanttracking der Vokale problemlos alle Formanten findet.

In Praat ist es möglich, Skripte zu erstellen, mit deren Hilfe diverse Prozessabfolgen zusammengefasst werden. Die Abfolge der Formantanalyse baut sich wie folgt auf:

- 1) Um die Formantwerte einer Audiodatei in Praat zu sammeln, müssen zuerst die Spektrogramm- und Formanteinstellungen angepasst werden um die Erkennung der Formantbereiche für jeden Sänger individuell zu optimieren.
- 2) Ist ein sauberes Formanttracking erreicht, wird ein Formantobjekt erstellt – eine Ebene, in der allein die Formantspur liegt.
- 3) Ein Texttraster wird über der Audiodatei erstellt. In diesem werden manuell die Zeitbereiche auftretender Vokale definiert, wobei Bereiche einer fehlerbehafteten Formantenerkennung ausgeschlossen werden können. Außerdem wird der Ein- und Ausschwingvorgang aus der Messung ausgeschlossen, da es an diesen Stellen häufig zu Problemen bei der Erkennung von Formanten kommt. Auch Konsonanten können mit dieser Methode von der Messung ausgeschlossen werden (siehe Abbildung 11).

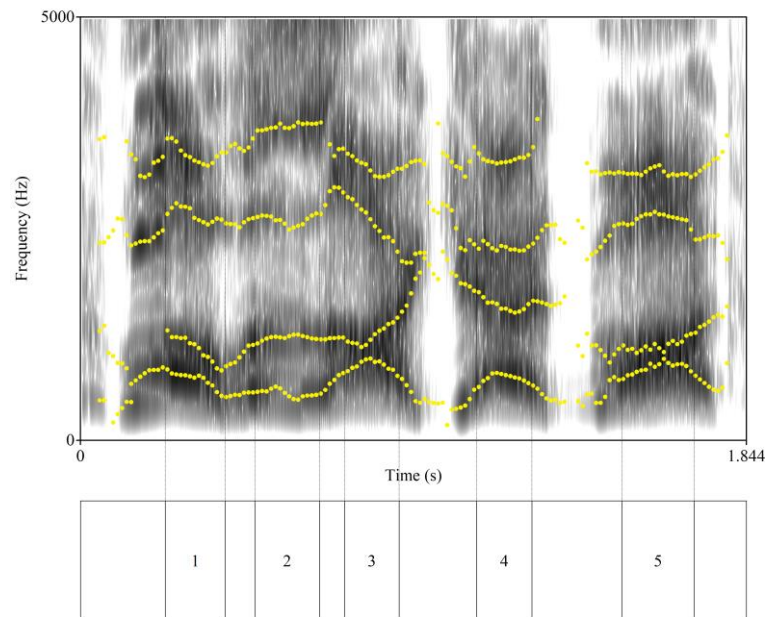


Abbildung 11: Ansicht der Formantspuren in Praat. Die Bereiche auftretender Vokale werden markiert.

Quelle: Eigene Darstellung

Im nächsten Schritt kommt ein Skript zum Einsatz (siehe Anlage 1), welches eine Auswahl der Formantwerte definiert, erfasst und in einer Tabelle auflistet. Der Prozess dieses Skripts verläuft wie folgt:

- 1) Es wird festgelegt, dass je definierten Zeitbereich 10 Zeitwerten ausgewählt werden. Dadurch kann die Varianz verschiedener Zeitbereiche eingeschränkt werden.
- 2) Praat greift auf das Textobjekt zu und erfasst einen manuell festgelegten Zeitbereich. Dabei werden nur gekennzeichnete Textfelder berücksichtigt.
- 3) Eine Tabelle wird erstellt, in welcher der Name des Zeitbereichs (1, 2, 3 usw.), dessen Zeitindex (Anfangs- und Endpunkt) und die Formanten F1 bis F4 aus dem Formantobjekt aufgelistet werden.
- 4) Der Prozess wird wiederholt, bis alle Zeitbereiche abgearbeitet sind.

Letztendlich wird die Tabelle als ein tab-separiertes Textdokument gespeichert und kann in dieser Form für weitere Schritte in Excel importiert werden.

4.2.2 Darstellung der Formanten im Vokaldreieck

Im zweiten Schritt wird die Lage der Formanten visualisiert. Um die Auswirkung der Vokalformung auf die Klangfarbe zu untersuchen, werden F1 und F2 in einem Vokaldreieck dargestellt. Im Durchschnitt können 17 bis 20 Zeitbereiche auftretender Vokale in einer Aufnahme erfasst werden, abhängig davon, wie viele Vokale zuverlässig getrackt wurden. Denn Praat zeigte teilweise Schwierigkeiten, die Genauigkeit der Formantenverfolgung bei atonalen Klängen wie Screaming zu gewährleisten.

Alle Formantentabellen wurden folglich in Excel importiert und die jeweiligen Frequenzen den Achsen eines Koordinatensystems zugeordnet. F1 entspricht der y-Achse und F2 der x-Achse. Dabei ist zu beachten, dass die Achsen eines Vokaldreiecks invertiert werden, sodass sich der Nullpunkt des Koordinatensystems in der oberen rechten Ecke der Darstellung befindet (siehe Abbildung 12).

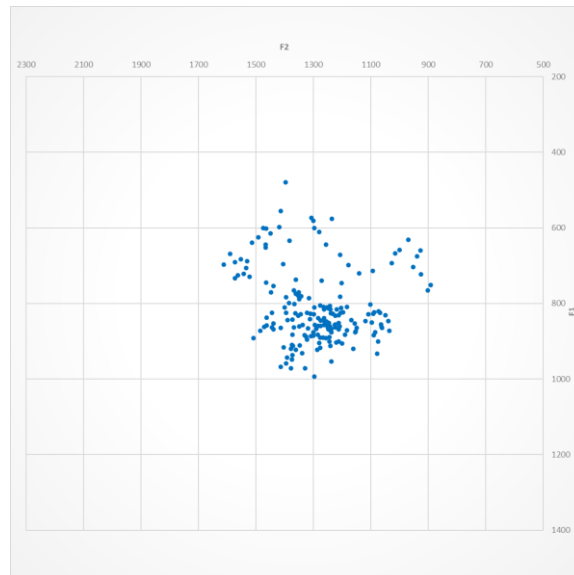


Abbildung 12: Darstellung von Formantmesspunkten im Vokaldreieck

Quelle: Eigene Darstellung

Um die verschiedenen Klangfarbenkategorien eines Sängers darzustellen, wird für jede Formantentabelle eine Datenreihe im Koordinatensystem angelegt. Jede Datenreihe erhält eine gesonderte Farbe, um die Datenpunkte jeder Reihe optisch differenzierbar zu machen. *FC Low* wird in grüner Farbe angezeigt, *FC Mid* blau, *FC High* rot. Die Achsen werden so formatiert, dass die Minimal- und Maximalwerte des Sängers, der diese erreicht, im Koordinatensystem sichtbar werden. Jedes Vokaldreieck erhält eine identische Formatierung. F1 wird zwischen 200Hz und 1,5kHz angezeigt, F2 zwischen 500Hz und 2,8kHz.

4.2.3 Darstellung der Formanten im Kastendiagramm

In einem Kastendiagramm können die wichtigsten Lagen und Streuungsmaße der ermittelten Formantdatenreihen visualisiert werden. Es ermöglicht auch, F3 und F4 in die Analyse mit einzubeziehen, da diese gewöhnlich nicht im Vokaldreieck dargestellt werden. Abbildung 15 zeigt ein Kastendiagramm der F1- bis F4-Datenreihen von S1 für die Kategorien *FC Low*, *FC Mid* und *FC High*. Der Median zeigt den Mittelpunkt der Verteilung der Datenmenge an und ist mit einem waagerechten Strich gekennzeichnet. Es liegen jeweils 50% der Daten oberhalb und 50% unterhalb des Medians. Quartile definieren die Mitte der Verteilung unterhalb des Medians (1. Quartil) und oberhalb des Medians (3. Quartil). Ein Quartil beinhaltet demnach 25% der Datenmenge. Der Interquartilsabstand (IQA) bildet sich zwischen 1. und 3. Quartil und ist als Kasten dargestellt (siehe Abbildung 13). Dieser ist eine Darstellungsform der Verteilung der Daten. Des Weiteren können einem Kastendiagramm Minimal- und Maximalwerte, sowie Ausreißer entnommen werden. Letztere sind Datenpunkte, die mehr als den 1,5-fachen Interquartilsabstand von Q1 bzw. Q3 aufweisen. Der Mittelwert der Datenreihe wird mit einem kleinen Kreuz gekennzeichnet.

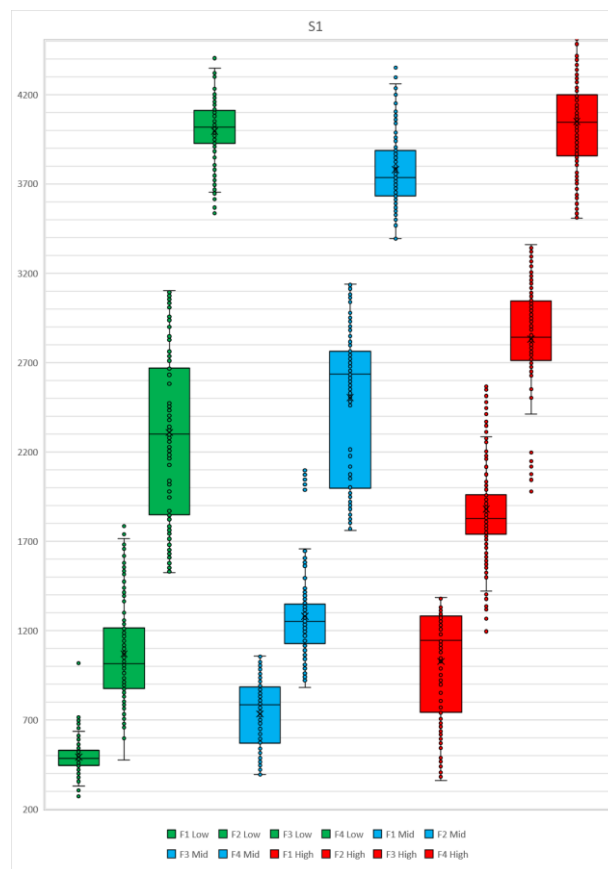


Abbildung 13: Darstellung der Formantmessreihen für verschiedene Klangfarben im Kastendiagramm
Quelle: Eigene Darstellung

Das Kastendiagramm hat allerdings den Nachteil, dass der Zusammenhang eines Messpunktes nicht nachvollziehbar ist, also welche Formantwerte einem gemeinsamen Messpunkt entspringen. Das Kastendiagramm wird daher hauptsächlich für die Einschätzung der Formantverteilung verwendet.

4.2.4 Darstellung des Amplitudenspektrums

Für den Klangvergleich der False Cord Screams werden die gemittelten Lautstärkepegel der Aufnahmen auf einen identischen Wert skaliert, um die Formantpegel unterschiedlicher Aufnahmen vergleichbar darzustellen zu können.

Das Ziel ist es, die Amplitudenspektren der drei Klangkategorien in einer gemeinsamen Grafik darzustellen. Praat bietet die Möglichkeit, erstellte Objekte in einem Fenster zeichnen zu lassen. Allerdings sind die Darstellungsoptionen begrenzt. Daher wird Praat lediglich benutzt, um Wertetabellen der erstellten Amplitudenspektren zu erzeugen. Für die Darstellung wird auf das Programm R zurückgegriffen, welches seinen Anwendungsbereich meist in analytischer und statistischer Erfassung und Darstellung von Daten findet.

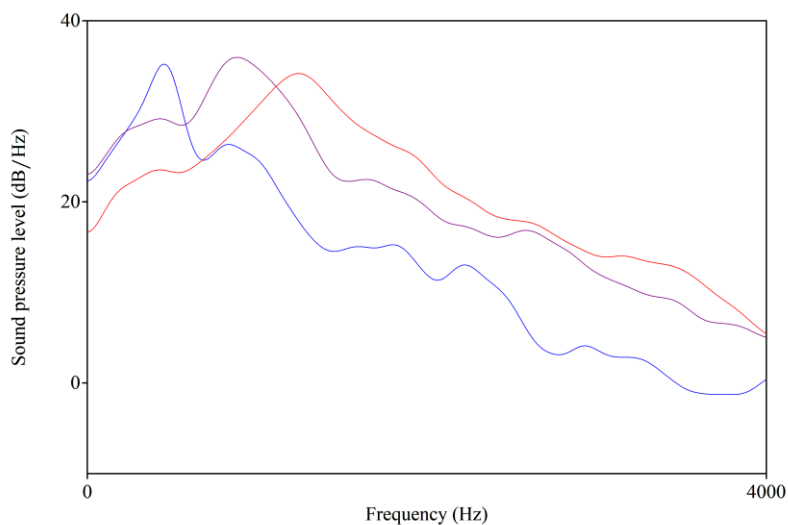


Abbildung 14: Darstellung der Amplitudenspektren von verschiedenen Klangfarben in Praat.

Quelle: Eigene Darstellung

Um in Praat eine Tabelle der Werte des Amplitudenspektrums zu erstellen, wird ein Skript verwendet, welches folgenden vereinfachten Prozessablauf beinhaltet:

- 1) Es wird eine manuelle Auswahl der zu analysierenden Audiodateien erfragt.
- 2) Für jede Audiodatei wird ein gleichnamiger Vektor erstellt, welchem später die Werte dieser Datei zugeordnet werden.
- 3) Für jede Audiodatei wird ein Spektrum erstellt.
- 4) Es wird die Tabelle definiert, in welche die Daten jeder Audiodatei ihrem gleichnamigen Vektor zugeschrieben werden.
- 5) Die erstellten Tabellen werden zusammengeführt und können als kommaseparierte Textdatei gespeichert werden.

Es folgt die Darstellung der Amplitudenspektren in dem Programm R. In einem Koordinatensystem kann nun für jeden Vektor die Amplitude zu ihrer zugehörigen Frequenz als Funktion dargestellt werden. Dazu wird *ggplot2* verwendet, welches ein System zur deklarativen Erstellung von Grafiken ist. Dieses wird ergänzt durch *zoo*, ein weiteres R-Paket, welches eine Infrastruktur der geordneten indizierten Beobachtungen bereitstellt. Diese wird intern in einem Vektor gespeichert und dient als Hilfsmittel zur Glättung der Amplitudenfunktionen.

Ein Skript in R führt einen Prozessablauf aus, welcher wie folgt vereinfacht erklärt werden kann:

- 1) R liest die *ggplot*- und *zoo*-Bibliotheken ein und ebenso die mit Praat erstellte Spektrumtabelle.
- 2) Die einzelnen Spektren werden erstellt. Dafür wird eine Anzeigetafel definiert, also ein Koordinatensystem mit den entsprechenden Achsen und einer passenden Skallierung. Um F1 bis F4 darstellen zu können wird der Frequenzverlauf von 0Hz bis etwa 4kHz dargestellt. Die Amplitude (*intensity*) wird zwischen 0dB und 45dB angezeigt. Anschließend wird jedes Spektrum mit zugehöriger Beschriftung auf seine Anzeigetafel projiziert.
- 3) Die Spektren und deren Beschriftungen werden auf einer gemeinsamen Anzeigetafel dargestellt.
- 4) Die Funktionen werden mit *zoo* geglättet. Der Glättungsfaktor beträgt dabei $k=3000$.

- 5) Die Funktion wird erneut mit der Glättung auf die Anzeigetafel projiziert.
- 6) Die Grafik wird gespeichert.

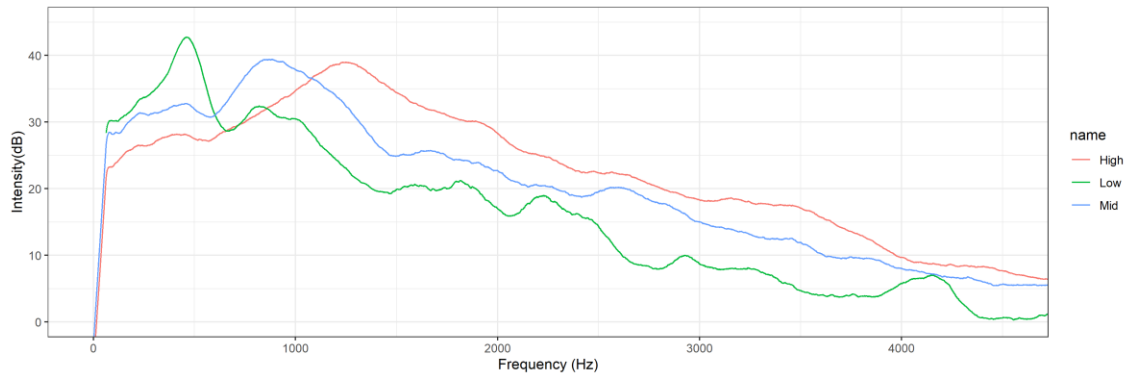


Abbildung 15: Darstellung der Amplitudenspektren verschiedener Klangfarben mit der Software R.
Quelle: Eigene Darstellung

Eine zusammengefasste Darstellung des Amplitudenverlaufs der verschiedenen Klangfarben ermöglicht nun einen genaueren Vergleich der Klangfarbe. Mit den zuvor gewonnenen Einblicken aus dem Vokaldreieck und Kastendiagramm über die Formantverteilungen, können nun auch Aussagen über die Formantpegel und Formantbandbreiten getroffen werden.

5 Ergebnisbericht

5.1 Klanganalyse

5.1.1 Formantanalyse S1

Die gesanglichen Darbietungen von S1 waren in der Kategorie *FC* dem Death-Metal-Genre zuzuordnen. Dabei wurde ein Titel in verschiedenen Klangfarben ausgeführt, was in diesem Fall eine sehr exakte Gegenüberstellung ermöglicht.

Klangfarbentendenz

Das Kastendiagramm veranschaulicht die Variation von F1 bis F4 in Abhängigkeit des gewählten Klangmodus. Eine grundlegende Auffälligkeit ist das Ansteigen derer Mittelwerte mit zunehmender Helligkeit der Klangfarbe. Im *FC Low* (siehe Abbildung 16, grün dargestellt) bildet sich der Mittelwert von $F1_{av}$ (average) bei etwa 500Hz aus, $F2_{av}$ bei etwa 1kHz. Ordnet man diese Werte gemäß der Tabelle der gemittelten Formantlagen des Vokaldreiecks ein, so tendiert der resultierende Vokalklang zu einem „o“. Es ist zu beachten, dass die Mittelwerte der Formanten nicht immer genau ihren tatsächlich Resonanzbereichen entsprechen, wie sie sich im Amplitudenspektrum äußern. Daher kann diese Methode nur eine Tendenz der Klangfarbe beschreiben. Für *FC Mid* (siehe Abbildung 16, blau dargestellt) steigt $F1_{av}$ auf 750Hz, $F2_{av}$ auf 1,3kHz. Dies entspricht eher einem „å“-Vokalklang, der etwas geschlossener ist als das herkömmliche „a“ und artikulatorisch näher an einem „o“ liegt. Im *FC High* (siehe Abbildung 16, rot dargestellt) stimmt die Frequenzpositionierung von $F1_{av}=1\text{kHz}$ und $F2_{av}=1,8\text{kHz}$ eher mit einem „ä“ überein.

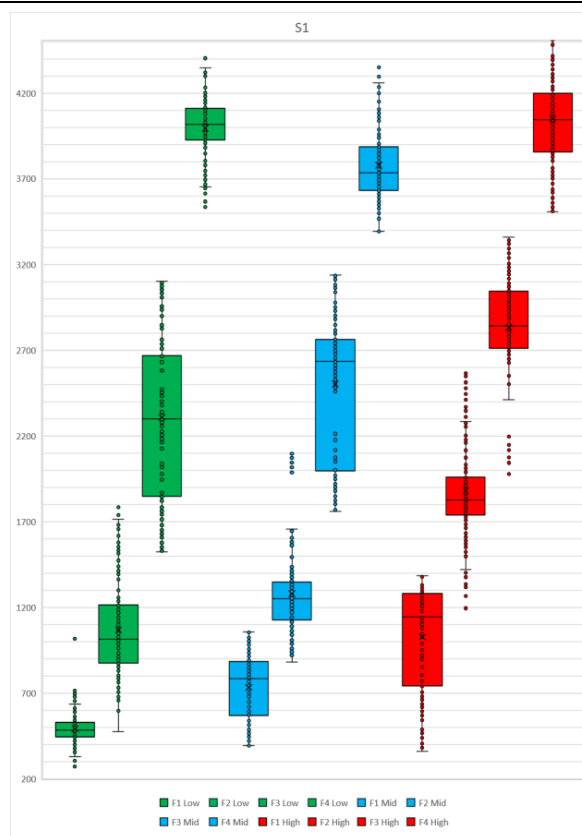


Abbildung 16: Kastendiagramm der F1- bis F4-Datenreihen von S1 für die Kategorien FC Low (grün), FC Mid (blau) und FC High (rot).
Quelle: Eigene Darstellung

Formantverteilung

Das Minimum ($F1_{\min}$) ändert seine Lage nur geringfügig, während das Maximum ($F1_{\max}$) zusammen mit dem entsprechenden Mittelwert einen linearen Anstieg verzeichnen. Folglich erweitert sich die Verteilung von F1 mit einer helleren Klangfarbe. Auch der Vergleich des Interquartilsabstands (IQA) stützt diese Beobachtung. Beginnend bei einem Wert von 100Hz im *FC Low* steigert sich der IQA über 300Hz im *FC Mid* hin auf 550Hz im *FC High*. Ein höherer F1 impliziert dabei eine offenere Vokalartikulation.

$F2_{\min}$ steigt mit höherem Modus an, ebenso wie der Mittelwert $F2_{av}$. Es ist zu bemerken, dass die Differenz zwischen $F1_{av}$ und $F2_{av}$ erst im *FC High* größer wird, was auf eine vermehrte Artikulation vorderer Vokale hindeutet.

Ebenso nimmt $F3_{av}$ mit fortschreitendem Modus zu, während die Breite der F3-Verteilung abnimmt. Gespreizte Lippen führen zu einem hohen F3 und werden typischerweise mit vorderen Vokalen in Verbindung gebracht. Dies deutet darauf hin, dass

in hellen Klangfarben sowohl die vorderen als auch die hinteren Vokale weiter vorn, also mit gespreizteren Lippen artikuliert werden.

$F4_{av}$ scheint nicht direkt vom Klangfarbenmodus beeinflusst zu werden, obwohl die Verteilung der erfassten Formantwerte einen leichten Anstieg zeigt. Mit zunehmendem Modus wird $F4$ im Amplitudenspektrum weniger erkennbar, da das Spektrum stärker gestreut wird.

Die Mittelwertbereiche der Formanten 1 und 2 sind im Vokaldreieck mit farbigen Kreisen markiert (siehe Abbildung 17). Bei *FC High* tritt der Resonanzbereich besonders gut hervor und spiegelt sich deutlich im Maximum von $F1$ wider.

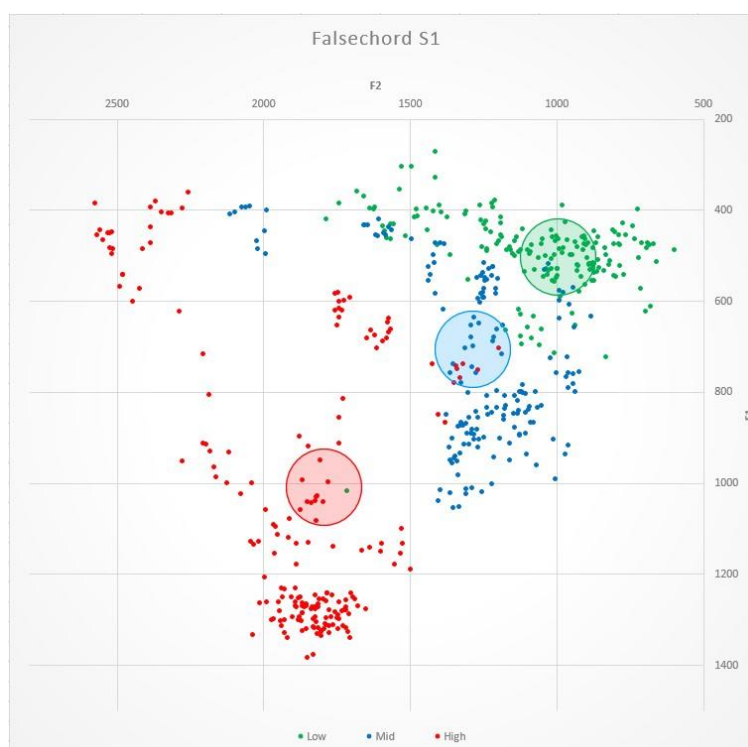


Abbildung 17: Darstellung von F1- und F2- Datenreihen der verschiedenen Klangfarben von S1 im Vokaldreieck. Die Mittelwerte sind mit farbigen Kreisen markiert.

Quelle: Eigene Darstellung

Andererseits können durch die Betrachtung der Amplituden von $F1$ und $F2$ deutlichere Rückschlüsse gezogen werden:

Formantpegel und Formantbandbreiten

Im Modus einer dunklen Klangfarbe ist F1 bei etwa 450Hz deutlich schmalbandiger und bildet den Bereich der größten Energie. Dieser Frequenzbereich ist, wie bereits erwähnt, auch in helleren Klängen der Stimme vorhanden, da $F1_{\min}$ nahezu in seinem Wert konstant bleibt. Diese Minimalwerte sind im *FC High* tatsächlich auf die vorderen Vokale zurückzuführen (Abbildung 18), bei welchen ohnehin F2 von größerer Relevanz für die artikulatorische Wahrnehmung ist. Hintere Vokale erfahren eine Anhebung der F1-Werte. Die Energie von $F1_{\min}$ nimmt demnach mit hellem Timbre drastisch ab und zeigt eine Reduzierung von -10dB im *FC Mid* (blau), und -15dB im *FC High* (rot). Je heller der Modus, desto höher und breitbandiger wird die Ausprägung von F1 in ihrer Frequenz in der Darstellung. Eine mögliche Erklärung hierfür ist die zunehmend breitere Verteilung der Formanten F1 und F2, da das Amplitudenspektrum den Mittelwert über die gesamten Aufnahmedauer wiedergibt.

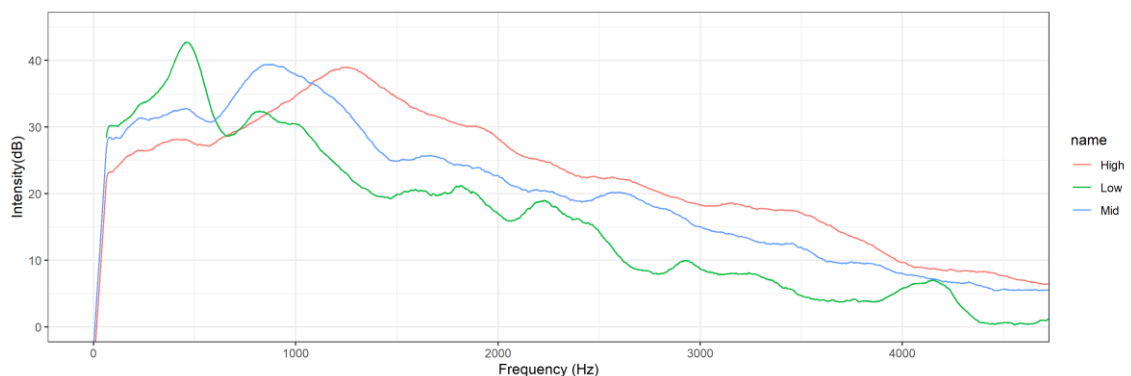


Abbildung 18: Vergleich der Amplitudenspektren der drei Klangfarben von S1.
Quelle: Eigene Darstellung

Somit führt eine breitere Verteilung zu einer größeren Bandbreite der Formantbereiche im Amplitudenspektrum. Diese Beobachtung gilt auch für höhere Formanten. Folglich wird es mit hellerem Klangmodus schwieriger, Formantbereiche mit dieser Methode im Amplitudenspektrum zu analysieren. $F2_{av}$ hält in jedem Modus einen relativ konstanten Pegel bei und zeichnet sich durch zunehmende Frequenz und Breitbandigkeit aus. Da die Formanten weitestgehend ihre Intensität behalten, im Frequenzspektrum jedoch verschoben werden, äußert sich diese Energieverschiebung im Amplitudenspektrum. Die Resonanzbereiche von F1 und F2 entsprechen den Bereichen der höchsten Messwertdichte im Vokaldreieck.

Die Klangfarbeneinstellungen im gleichen musikalischen Kontext von S1 machen eine zeitlich exakte Übereinstimmung der Sonagramme möglich (Abbildung 19). Es zeigt sich, dass für *FC Low* (oben) die Formantbereiche sehr deutlich erkennbar sind. Mit höherer Klangfarbe offenbart sich eine zunehmende Diffusität der Bereiche größter Energie. Des Weiteren ist zu erkennen, dass dabei die Energieverteilung deutlich angehoben wird.

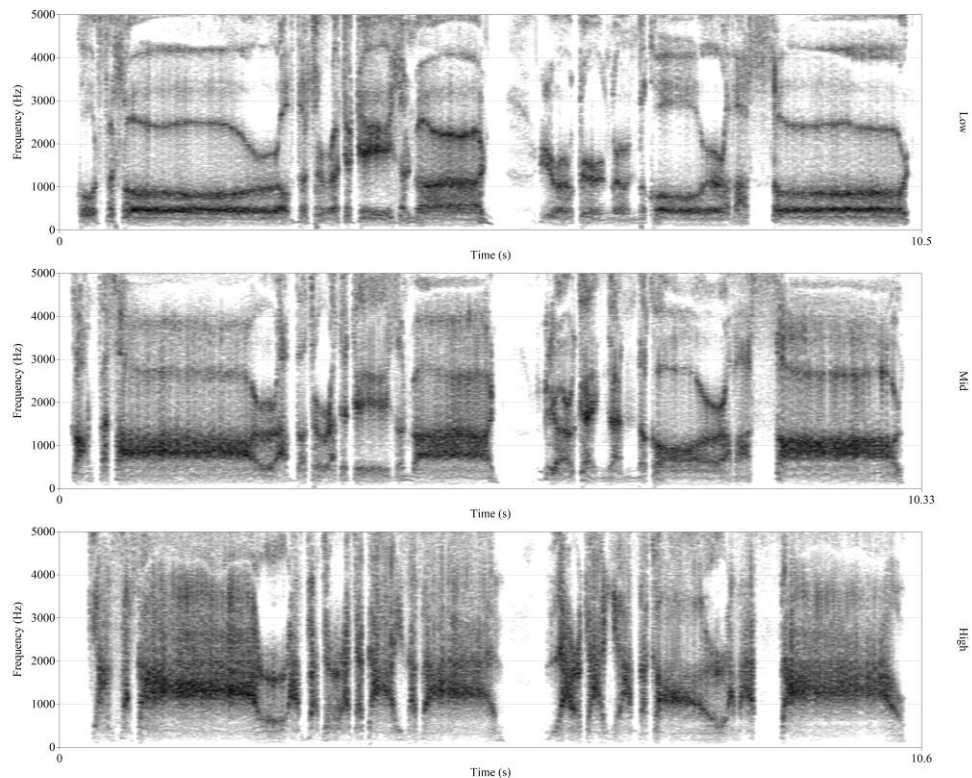


Abbildung 19: Darstellung der drei Klangfarben von S1 im Breitband-Sonagramm.
Quelle: Eigene Darstellung

Fazit

Die Beeinflussung der Klangfarbe kann auf die simultane Erhöhung von F1 und F2 zurückgeführt werden, wobei eine gleichzeitige Erhöhung der Differenz zwischen F1 und F2 eine hellere Klangfarbe durch markante vordere Vokale zusätzlich verstärkt. Gleichzeitig steigt mit höherem Klangfarbenmodus die Verteilung des ersten Formanten. Die Verteilung des zweiten Formanten bleibt nahezu konstant und wird in ihrer Gesamtheit im Frequenzbereich angehoben.

Für die Vokalartikulation bedeutet dies folgendes:

- 1) Offene Vokale werden im *FC Low* sehr geschlossen und hinten artikuliert (niedriger F1 und F2).
- 2) Offene Vokale werden im *FC High* mit sehr gespreizten Lippen (hoher F3), offen und deutlich weiter vorn artikuliert (kleinerer Mundraum).
- 3) Geschlossene Vokale bilden im *FC Low* die allgemeine Klangtendenz. Geschlossene vordere Vokale werden dabei in ihrer Ausprägung limitiert.
- 4) Vordere Vokale sind im *FC High* ähnlich stark geschlossen wie im *FC Low*, werden im *FC High* allerdings so weit vorn wie möglich artikuliert. Der Mundraum wird dabei sehr klein geformt und ein hoher F2 erzielt.
- 5) Die Formung hinterer Vokale wird im *FC Low* bevorzugt. Im *FC High* werden hintere Vokal in ihrer Tendenz mehr zu vorderen Vokalen geformt.

5.1.2 Formantanalyse S2

Sänger 2 war ebenfalls in der Lage, False Cord Screams in den drei kategorisierten Klangfarben im musikalischen Kontext der Genre Metalcore bzw. Progressive Metal zu erzeugen. Bei der Untersuchung der Mittelwerte der ersten Formanten lassen sich, ähnlich wie bei der Analyse von S1, folgende klangliche Tendenzen beobachten:

Im *FC Low* wird ein „o“ bzw. „å“-artiger Klang erreicht. $F1_{av}$ befindet sich bei 650Hz, $F2_{av}$ bei 1,1kHz. Im Modus *FC Mid* wird $F1_{av}$ bei 850Hz, $F2_{av}$ bei 1,3kHz erzeugt, was zu einem Klang führt, der charakteristisch mit dem Vokal „a“ vergleichbar ist. Ein Unterschied zu S1 ist zu beobachten, wenn eine helle Klangfarbe angestrebt wird. $F1_{av}$ bleibt in seiner Lage bestehen, nur $F2_{av}$ erhöht sich weiter auf 1,75kHz. Die Klassifizierung der Formantwerte weist in diesem Fall eine Tendenz zum Vokal „ä“ auf.

Es zeigt sich, dass $F1_{\min}$ mit steigendem Modus nur minimale Veränderungen in seiner Lage aufweist. Die Wahrnehmung der Klangfarbe kann einerseits auf die simultane Anhebung von $F1_{av}$ und $F2_{av}$ (*FC Low* zu *FC Mid*), andererseits auf die Anhebung von $F2_{av}$ bei konstantem $F1_{av}$ und derer resultierenden erhöhten Differenz zurückgeführt werden (*FC Mid* zu *FC High*) (Abbildung 20 und Abbildung 21). Durch letzteres Phänomen kann die Tendenz zu einer vorderen Artikulation geschlussfolgert werden.

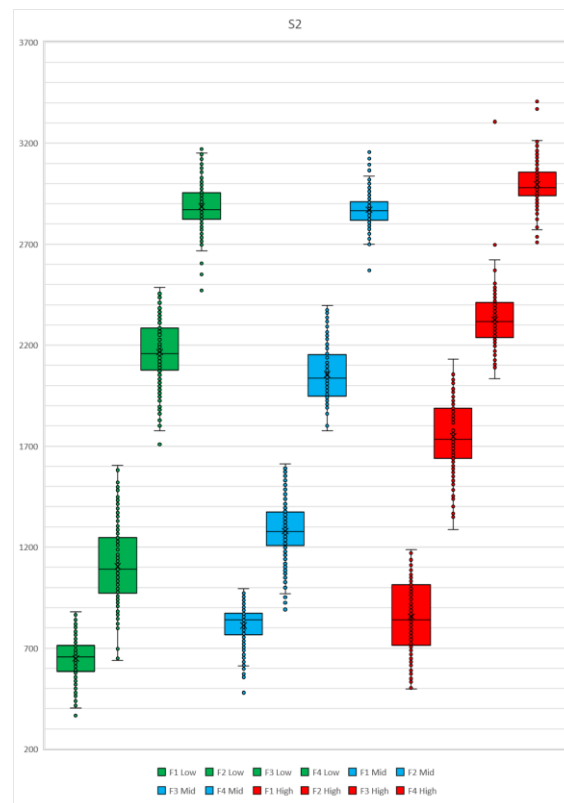


Abbildung 20: Kastendiagramm der F1- bis F4-Datenreihen von S2 für die Kategorien *FC Low* (grün), *FC Mid* (blau) und *FC High* (rot).
Quelle: Eigene Darstellung

$F3_{av}$ weist ebenso wie $F4_{av}$ nur geringfügige Frequenzverschiebungen auf. Es lässt sich vermuten, dass F3 und F4 im *FC High* von der Differenz zwischen F1 und F2 beeinflusst werden (Abbildung 20). Letztlich wird die Steuerung der Klangfarbe in jedem Modus hauptsächlich durch die Manipulation von F1 und F2 erreicht. Als Entwicklung klanglicher Eigenschaften lässt sich der Übergang von einer geschlossenen hinteren Artikulation (*FC Low*) zu einer offenen vorderen Artikulation (*FC High*) ableiten.

- 1) Offene Vokale werden für eine helle Klangfarbe sowohl extremer geöffnet, als auch weiter vorn artikuliert (steigender F2).
- 2) Geschlossene hintere Vokale werden dabei offener und weiter vorn artikuliert.

- 3) Im *FC High* werden hintere Vokal in ihrer Tendenz mehr zu vorderen Vokalen geformt und kommen geschlossen („u“ oder „o“) weniger häufig vor und werden in ihrer Ausprägung limitiert (Abbildung 21).

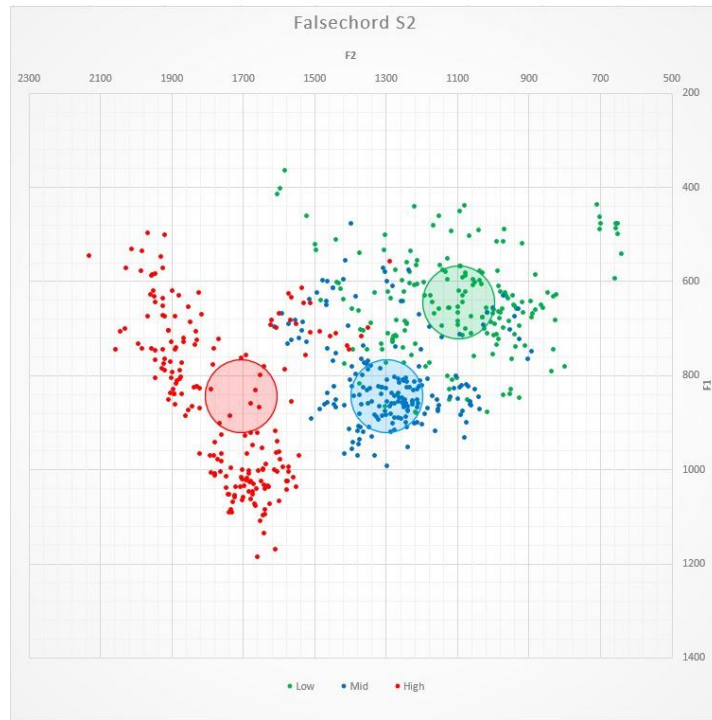


Abbildung 21: Darstellung von F1- und F2- Datenreihen der verschiedenen Klangfarben von S2 im Vokaldreieck. Die Mittelwerte sind mit farbigen Kreisen markiert.

Quelle: Eigene Darstellung

Weitere Erkenntnisse liefert das Amplitudenspektrum. Während bei *FC Low* F1 betont wird und einen deutlichen Pegelunterschied zu F2 von etwa 10dB aufweist, nähern sich die Pegel bei *FC Mid* nahezu vollständig an. Dieser Zusammenhang bleibt auch bei *FC High* erhalten, wobei die simultane Anhebung der Frequenz von F2 deutlich wird (Abbildung 21). Dies äußert sich im Frequenzbereich von 1,6kHz bis 2kHz.

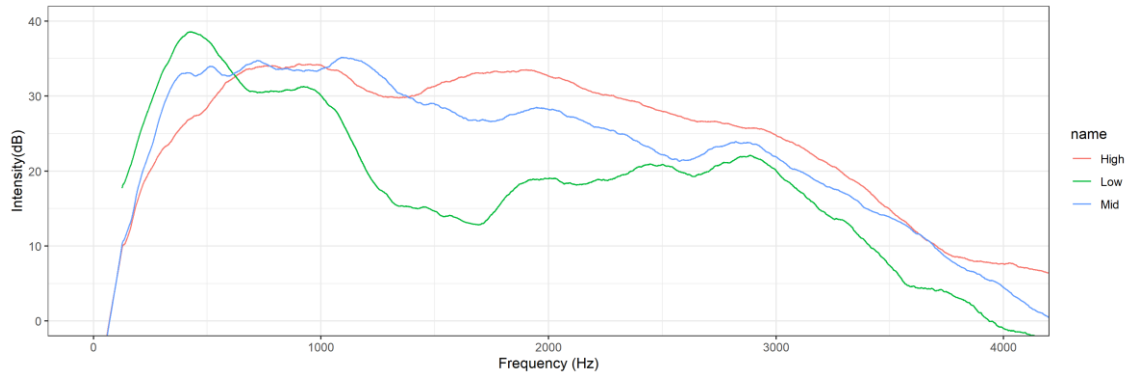


Abbildung 22: Vergleich der Amplitudenspektren der drei Klangfarben von S2.
Quelle: Eigene Darstellung

5.1.3 Formantanalyse S3

Sänger 3 deckte mit seinen gesanglichen Darbietungen die Kategorien *FC Low* und *FC Mid* ab, sodass *FC High* im Vergleich herausfällt. Die dargebotene Musikrichtung ist im Black Metal angesiedelt, wo normalerweise eher helle Klangfarben vorherrschen.

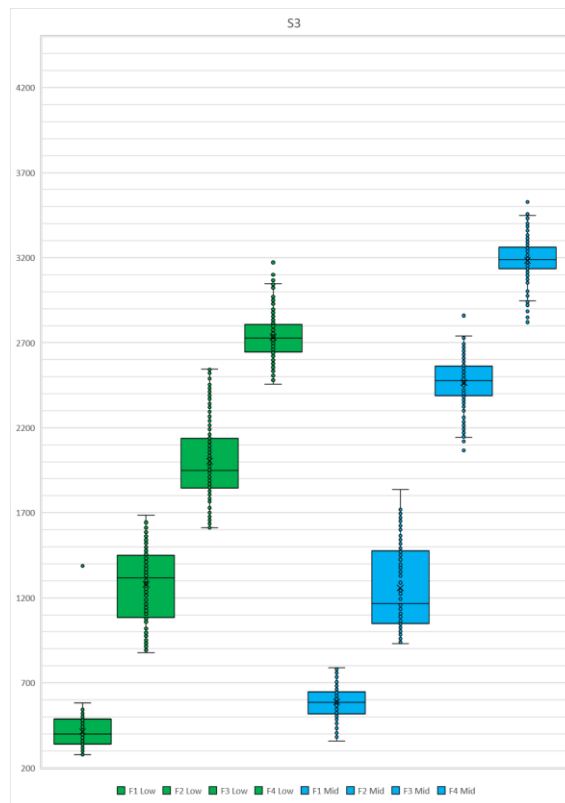


Abbildung 23: Kastendiagramm der F1- bis F4-Datenreihen von S3 für die Kategorien *FC Low* (grün) und *FC Mid* (blau).
Quelle: Eigene Darstellung

Im *FC Low* liegen $F1_{av}$ und $F2_{av}$ bei 400Hz bzw. 1,3kHz. Während $F1_{av}$ einen relativ niedrigen Wert ausbildet, ist eine große Differenz zwischen $F1_{av}$ zu $F2_{av}$ beobachtbar (siehe Abbildung 23). Zu berücksichtigen ist, dass S3 einen großen Körperbau aufwies, was anatomisch auf einen größeren Vokaltrakt schließen lässt und einen niedrigen $F1_{av}$ erklären könnte. Anhand der gemittelten Formantlagen lässt sich eine Tendenz zum Vokal „u“ identifizieren. $F2_{av}$ entspricht in etwa einem „a“. Daraus kann geschlossen werden, dass der Sänger mit einem großem Rachenraum und tiefen ersten Formanten die Vokale relativ weit öffnet. Mit der erwarteten Erhöhung von $F1_{av}$ für *FC Mid* bleibt $F2_{av}$ nahezu konstant (siehe Abbildung 24). Bei $F1_{av}=600$ Hz und $F2_{av}=1,2$ kHz zeigt sich hier eine „o“- bzw. „å“-artige Vokaltendenz.

Die grundlegende Komponente der Klangfarbenanpassung ist in diesem Fall die Lage von F1. Es handelt sich in beiden Modi um ähnlich offene Vokale, wobei der Rachenraum für *FC Mid* vermutlich verengt wird, um einen höheren F1 zu erzeugen. Zusätzlich ist ein Anstieg der Formanten 3 und 4 zu erkennen (Abbildung 23). Die Lage von F3 weist auf weit gespreizte Lippen hin, was auch mit einem offenen Vokal übereinstimmt.

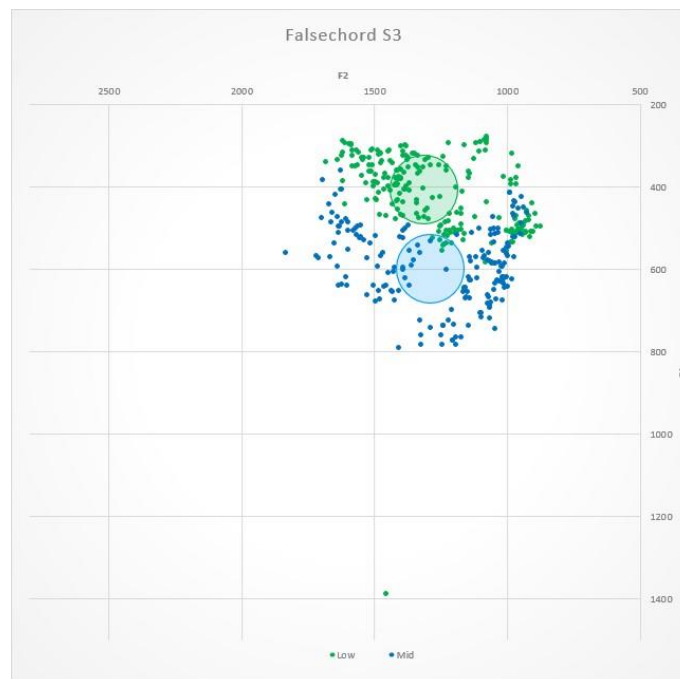


Abbildung 24: Darstellung von F1- und F2- Datenreihen der verschiedenen Klangfarben von S3 im Vokaldreieck. Die Mittelwerte sind mit farbigen Kreisen markiert.

Quelle: Eigene Darstellung

Das Amplitudenverhalten der Formanten zeigt, dass sich F1 primär bei 300Hz (nahe $F1_{\min}$), sekundär bei 500Hz ausprägt (siehe Abbildung 25). Daraus lässt sich ein $F1_{av}$ von 400Hz ableiten, obwohl an dieser spezifischen Frequenzlage keine F1 Resonanz vorliegt. Für *FC Mid* bestätigt $F1_{av}$ jedoch die tatsächliche Lage des stärksten Resonanzbereichs. E zeigt sich, dass für *FC Mid* F1 im Wesentlichen in seinem Minimum bestehen bleibt, allerdings an Pegel verliert und hingegen $F1_{av}$ dominiert. Geschlossene Vokale werden folglich bei hellerer Klangfarbe weniger in ihrem dunklen Klang betont. Zusätzlich macht sich für *FC Mid* ein Pegelanstieg des zweiten Formanten um etwa 5dB bemerkbar, welcher die Energieverteilung im Spektrum beeinflusst und ein helles Timbre bekräftigt. Außerdem kann davon ausgegangen werden, dass in einem dunklen Timbre offene Vokale etwas geschlossener sind, während in einem hellen Timbre geschlossene Vokale etwas offener artikuliert werden. F3 und F4 behalten von *FC Low* zu *FC Mid* eine ähnliche Intensität und verschieben sich mit konstantem Abstand zueinander in ihrer Frequenzlage nach oben.

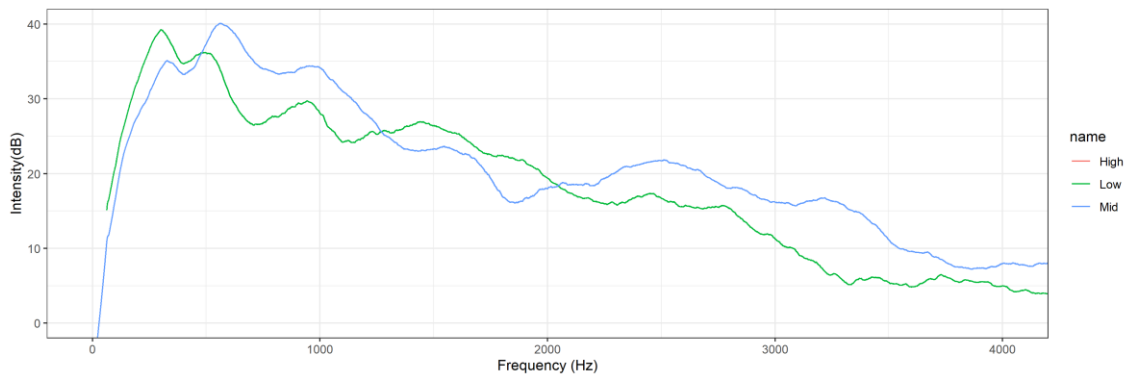


Abbildung 25: Vergleich der Amplitudenspektren Klangfarben *FC Low* und *FC Mid* von S3.
Quelle: Eigene Darstellung

5.1.4 Formantanalyse S4

Für Sänger 4 konnten ausschließlich Daten in der Kategorie *FC Mid* erhoben werden. Somit ist ein Klangfarbenvergleich nicht möglich. Dennoch kann die Klangfarbe mit den bereits ermittelten Tendenzen der anderen Sänger verglichen werden. Das Genre lässt sich dabei in den musikalischen Kontext des Progressive Metal einordnen.

Die Mittelwerte der Formanten zeigen eine vokale Tendenz zu einem „ö“ bzw. „ä“ auf. $F1_{av}$ prägt sich bei 500Hz aus, $F2_{av}$ bei 1,6kHz. Durch die Vokaltendenzen lässt sich auf eine starke Manipulation der Zungenposition schließen, um einen hellen Klang anzustreben. Es zeigt sich eine große Differenz zwischen ersten und zweiten Formanten, wie dies bereits bei S2 zu beobachten war. Vordere Vokale scheinen bevorzugt zu werden. Dies zeigt sich durch die Bündelung von Messdaten bei $F1=400\text{Hz}$ und $F2=2\text{kHz}$ (Abbildung 26).

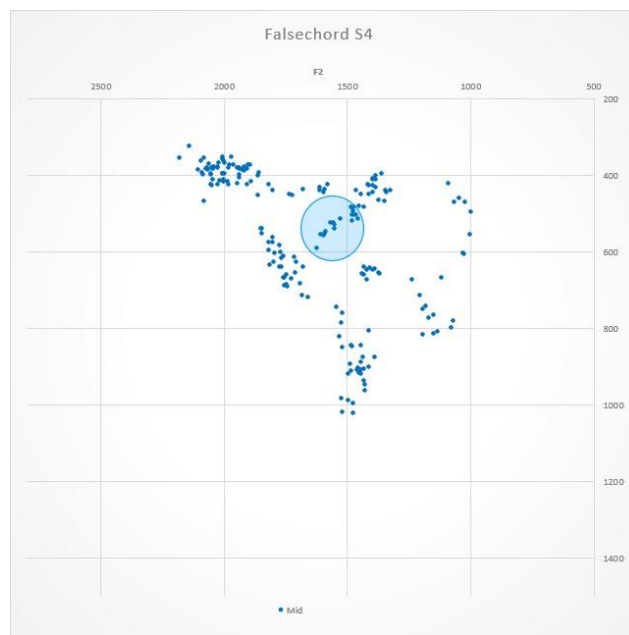


Abbildung 26: Darstellung von $F1$ - und $F2$ - Datenreihen der Klangfarbe $FC\ Mid$ von S4 im Vokaldreieck. Der Mittelwert ist mit einem farbigen Kreis markiert.
Quelle: Eigene Darstellung

Ein Blick auf das Amplitudenspektrum offenbart, dass sich die Verteilung von $F1$ relativ weit streut, was eine Einordnung in die Kategorie $FC\ Mid$ bekräftigt. Der Übergang von den $F1$ - und $F2$ -Frequenzen liegt nach Aussage des Vokaldreiecks bei etwa 1kHz. Darüber äußert sich $F1$ im Bereich von 1kHz bis etwa 2kHz.

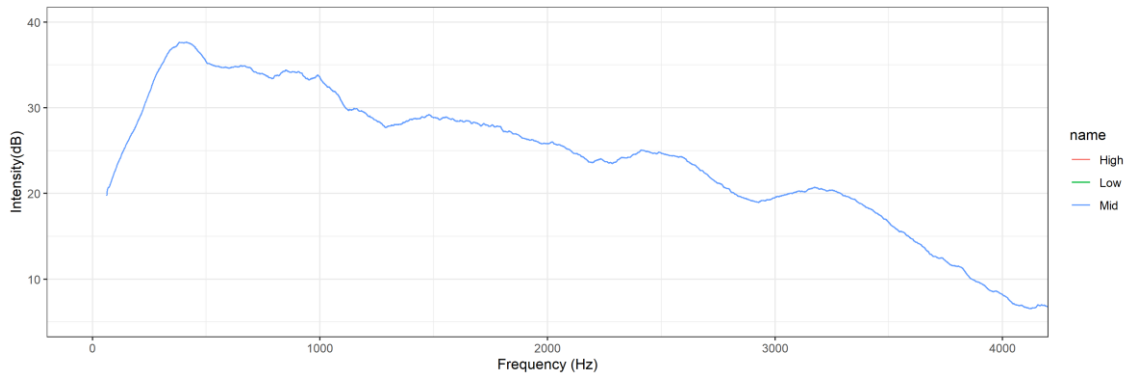
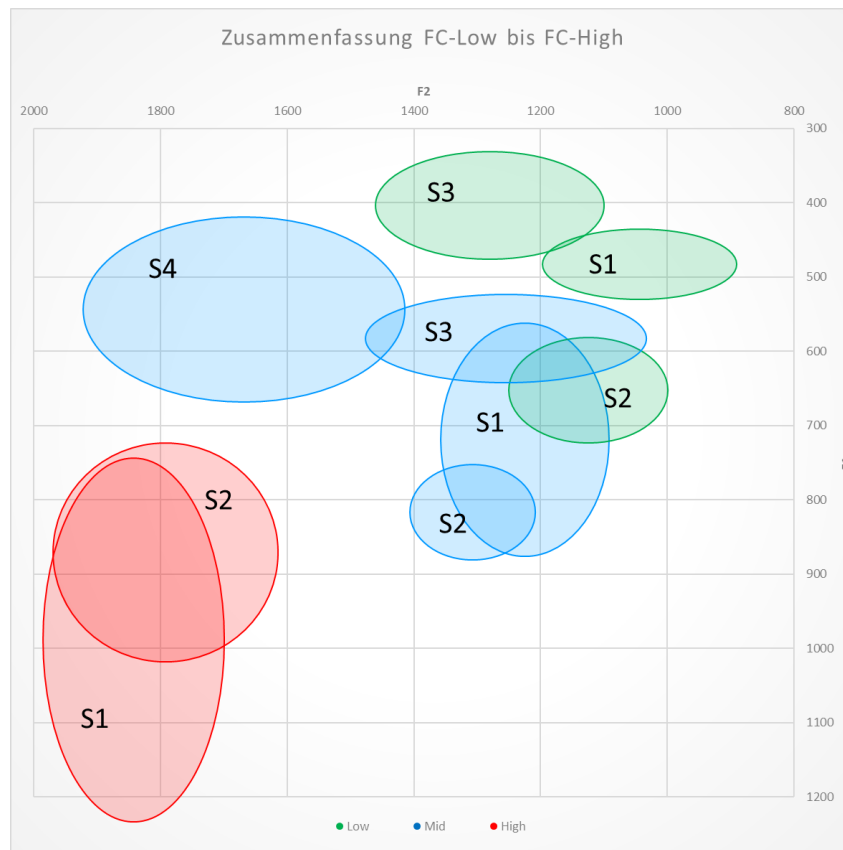


Abbildung 27: Darstellung des Amplitudenspektrums der Klangfarbe FC Mid von S4.
Quelle: Eigene Darstellung

5.2 Fazit

In der folgenden Abbildung (Abbildung 28) wird die Verteilung der Formantwerte von F1 und F2 zusammenfassend dargestellt. Die Höhe einer Ellipse entspricht dem IQA des ersten Formanten, die Breite dem IQA des zweiten Formanten. Dadurch kann die Verteilung der Formantlagen in der Visualisierung berücksichtigt werden, um eine Vorstellung der Klangeigenschaften auf Messdatenebene zu bekommen.

F1 scheint den größten Einfluss auf die Ausprägung der Klangfarbe zu haben, da dieser Formantwert in den Analysen durchweg direkt mit der „Höhe/Tiefe“ des Screams in Verbindung gebracht werden konnte. Niedrige F1-Werte mit einer stark komprimierten Verteilung führen zu einem dunklen Timbre. Dies sorgt für eine schmalbandige Charakteristik von F1 im Amplitudenspektrum, wobei die Energieverteilung eng konzentriert ist, um einen tiefen Wahrnehmungseindruck zu erzeugen. Die tiefste Resonanzfrequenz wird durch einen erheblichen Pegelunterschied zwischen F1 und F2 betont. Bei dunklen Klangfarben lässt sich auf eine überwiegend geschlossene Artikulation schließen. Offene Vokale werden in dunklen Klangfarben stark beeinträchtigt. Wobei beispielweise die Wörter „my life“ im englischen Gesang sehr offen sind, so würden sie in einer möglichst tiefen Klangfarbe sehr manipuliert erscheinen und eher als ein geschlossenes „moü loüf“ artikuliert werden. Dabei erzeugen gerundete Lippen den Klangunterschied vom „i“ zum „ü“ und betonen ebenfalls eine dunkle Klangfarbe.



*Abbildung 28: Darstellung der Mittelwerte verschiedener Klangfarben der Sänger unter Berücksichtigung der Formantenverteilung
Quelle: Eigene Darstellung*

Umgekehrt wird durch Anheben von F1 ein helles Timbre erreicht. Die Verteilung weitet sich aus, was auf eine vielfältigere Stimmgestaltung hinweist. Darüber hinaus können Anzeichen einer Art Formantenstimmung beobachtet werden, da F1 eine übermäßige Erhöhung aufweist. Hohe F1-Werte schließen außerdem auf eine überwiegend offene Artikulation. Ein gesungenes „not you“, welches überwiegend geschlossen artikuliert wird, erfährt in einem übermäßig hellen Timbre eine Vokalmanipulation zu einem „nat yao“. Die Manifestation von F1 ist in erster Linie vom Resonanzraum des Rachenraums abhängig und kann auf Faktoren wie die Kehlkopfposition, die Gaumenposition oder Zungenposition zurückgeführt werden.

F2 ist ebenfalls ein wichtiger Faktor für die Wahrnehmung dunkler bis heller Stimmklangfarben. Auch wenn die Beeinflussung der Klangfarbe in einem bestimmten Fall nicht auf eine Frequenzlagenänderung des zweiten Formanten zurückzuführen ist, steigt dieser bei hellem Timbre in seinem Pegel markant an. In anderen Fällen, die durch ein helleres Timbre gekennzeichnet sind, kommt es zu einer gleichzeitigen Erhöhung von F1 und F2. Diese können auf Anzeichen einer Formantenstimmung hin-

weisen und stehen im Zusammenhang mit dem Öffnen des Mundes. Die Verteilung von F2 weist im Vergleich zu F1 im Allgemeinen weniger ausgeprägte Merkmale auf.

Die Differenz zwischen F1 und F2 zeigt, dass die vorderen Vokale in ihrer Ausprägung bei der Erzeugung einer tiefen Klangfarbe stark begrenzt werden. „This is“ wird in einer dunklen Klangfarbe wohl eher als ein „thüs üs“ artikuliert. Dabei spielt wieder die starke Rundung der Lippen eine Rolle. Umgekehrt zeichnen sich Vordervokale in hellem Timbre durch einen großen Unterschied in ihren ersten Formanten aus. Eine Koordination, die dieses Formantenverhalten verstärken kann, ist Twang, bei welchem es zu einer Verengung des Kehltrichters kommt. Grundsätzlich kann eine Vergrößerung der Differenz zwischen F1 und F2 die wahrgenommene Helligkeit der Klangfarbe betonen.

Es zeigte sich, dass F3 überwiegend an den Anstieg der Formanten F1 und F2 gekoppelt ist. In einem bestimmten Fall von S2 (*FC Low* zu *FC High*) trat dies allerdings nicht ein.

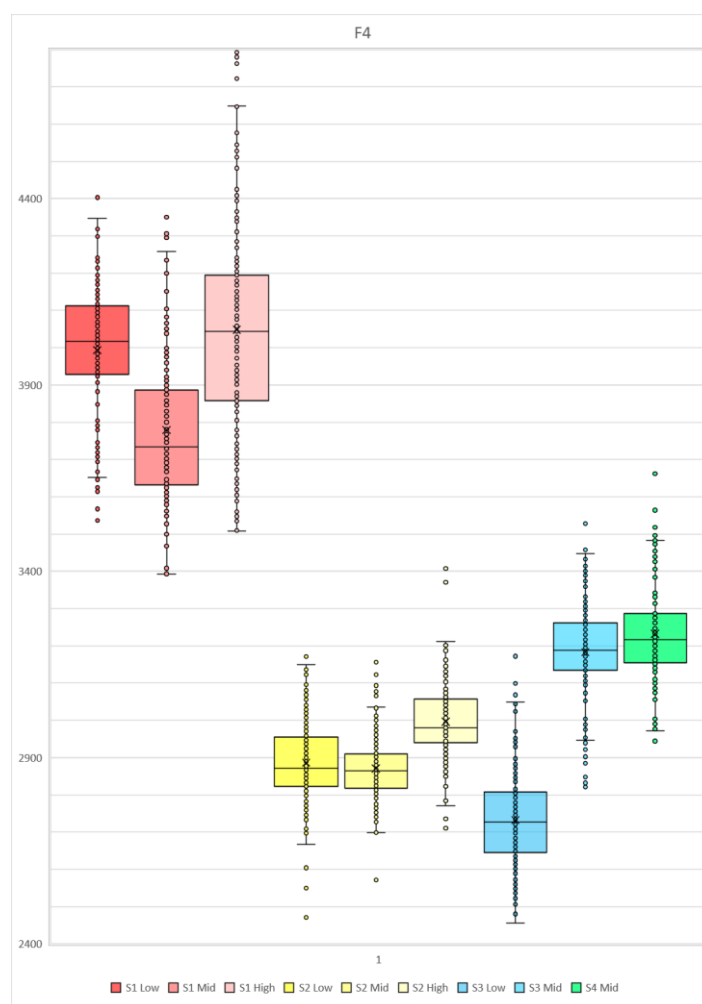


Abbildung 29: Vergleichende Darstellung der F4-Verteilungen der Sänger
Quelle: Eigene Darstellung

F4 zeigte sehr unterschiedliche Lageveränderungen (Abbildung 29). In der Regel veränderte sich diese Lage nur minimal und war weniger von dem Verhalten der Formanten F1 und F2 beeinflusst. Daher kann davon ausgegangen werden, dass F4 keinen entscheidenden Einfluss auf die Wahrnehmung der Klangfarbe hat und sich außerdem sehr sängerspezifisch ausprägt. Denn die Formanten F1 bis F3 prägten sich für alle Sänger in einem ähnlichen Frequenzbereich aus, wobei für F4 auffällige Unterschiede auftraten.

Die Betrachtung der Klangfarben im Kontext des Genres zeigt, dass im Death Metal der Stimmklang extrem dunkel gefärbt wird, um einen brachialen und bössartigen Klangeindruck zu erzeugen. Im Metalcore und anderen Subgenres wird hingegen auf ein breiteres Klangspektrum abgezielt, wobei diverse Stimmfarben als Effekte eingesetzt werden. *High Screams* erzeugen dabei eine markante und umfangreiche Vokalartikulation, um die Energiekonzentration der Stimme frequenziell anzuheben. Dies kann nützlich sein, damit sich das Screaming gegen komplexe und stark verzerrte Gitarrenparts durchsetzen kann. Ein übermäßig helles Timbre betont auch die Aggressivität und Durchsetzungskraft des Klangeindrucks eines False Cord Screams.

5.3 Optimierungsansätze

Im Laufe der methodischen Durchführung dieser Arbeit wurde klar, dass die Länge der musikalischen Darbietung und insbesondere die unterschiedliche Wortwahl verschiedener gesungenen Texte Einfluss auf die Messdaten der Formant- und Spektralanalyse und deren Vergleichbarkeit haben.

Eine isolierte Erfassung der Klangfarbeneinstellungen verschiedener Sänger zu erzielen und gleichzeitig ein vollständigeres Bild derer Vokalformung zu erfassen bedeutet, dass alle Vokalvariationen in die Auswertung einfließen. Um ein umfangreiches Vokaldreieck eines Sängers zu erstellen, müssen von diesem also je Klangmodus eine Reihe von Wörtern erbracht werden, welche alle artikulatorischen Variationen eines Vokals nach IPA enthalten. Jeder Vokal kann dann differenziert betrachtet und im Vokaldreieck dargestellt werden (Winn, Make your own vowel chart!, 2021). Letzten Endes können eindeutige Aussagen über die Formung jeden Vokals getroffen werden. Um den musikalischen Kontext herzustellen, wäre eine Überlegung, eine einheitliche Textpassage von jedem Sänger einsingen zu lassen, die alle Vokalvariationen einer Sprache enthält. Dieses Prozedere ist für verschiedene Klangfarben durchzuführen und bietet letztlich einen vergleichbareren Versuchsaufbau, angenommen es geht primär um eine rein technische Vergleichbarkeit. Auch eine spektrale Darstellung im Amplitudenspektrum kann in der Genauigkeit davon profitieren.

5.4 Schlusswort

Zusammenfassend hat diese Arbeit den Bereich der Klangfarbenmanipulation für Extreme-Metal-Gesang durch die Anwendung von Formantanalysetechniken untersucht. Durch die Auseinandersetzung mit der komplexen Natur der Stimmklangerzeugung und der Manipulation von Formanten wurden wertvolle Erkenntnisse über die nuancierten Eigenschaften und Ausdrucksmöglichkeiten extremer Metal-Gesänge gewonnen. Durch die umfassende Untersuchung und Analyse, die in dieser Arbeit vorgestellt wird, wird deutlich, dass Formantanalysetechniken ein leistungsstarkes Werkzeug zum Verständnis und zur Manipulation der Klangfarbe dieses einzigartigen Gesangsstils darstellen. Die hier präsentierten Erkenntnisse tragen zum breiteren Verständnis der Gesangstechniken in der extremen Metal-Musik bei und eröffnen neue Wege für Musiker und Toningenieure, um ihren kreativen Ausdruck zu verbessern und die Grenzen der klanglichen Möglichkeiten im Bereich der extremen Metal-Musik weiter zu verschieben. Im weiteren Verlauf der Forschung besteht die Hoffnung, dass diese Arbeit zukünftige Untersuchungen und Fortschritte im Bereich der Klangfarbenmanipulation inspirieren und letztendlich den Bereich des extremen Metal-Gesangs und das breitere Feld der Stimmklangerzeugung bereichern wird.

Literaturverzeichnis

Bernstein, E. (2019). *Elektroakustik*. München: Springer Vieweg.

Caffier, P. P. (2018). Common Vocal Effects and Partial Glottal Vibration in Professional Nonclassical Singers. *Journal of Voice*, 32, S. 340-346.

Dickreiter, M. D. (2013). *Handbuch der Tonstudioteknik* (8. Ausg.). Berlin: De Gruyter Saur.

Humphrey, E. R. (2019, Januar). An Introduction to Signal Processing for Singing-Voice Analysis. (I. S. Magazine, Ed.) pp. 82-94.

Kazi R., T. S.-E. (2004). A short biography on the life of the dedicated anatomist – Valsalva. *Journal of Postgraduate Medicine*, 314-315.

Lehmann, C. (18. 10 2020). *Christian Lehmann*. Abgerufen am 23. 04 2023 von https://www.christianlehmann.eu/ling/lg_system/phon/index.html?https://www.christianlehmann.eu/ling/lg_system/phon/05_Vokale.html

Lewis, J. W. (2023, 03 11). Retrieved 04 14, 2023, from SWPhonetics: <https://swphonetics.com/>

Linke, T. (10. März 2023). Metal Screaming Techniques. (A. Geier, Interviewer)

Maurer, D. (2016). *Acoustics of the Vowel*. Bern: Peter Lang.

McDonald Klimek, M. O. (2009). *Estill Voice Training Level One. Figures for voice control*. Estill Voice International.

McGlashan, J. &. (2008). *A study of extreme singing vocal effects including*. London, UK: British Voice Association.

Mesiä, S. &. (2015). Heavy Metal Vocals. A terminology compendium. *MODERN HEAVY METAL: MARKETS, PRACTICES AND CULTURES*, 383-392.

- Meyer, J. (1972). *Akustik und musikalische Aufführungspraxis* (3. Ausg.). Frankfurt am Main: Erwin Bochinsky.
- Ribaldini, P. (2019, April). Heavy metal vocal technique terminology compendium. University of Helsinki.
- Richter, B. E. (2017). *Die Stimme. Einblicke in die physiologischen Vorgänge beim Singen und Sprechen*. ROM-DVD: Helbling.
- Roubeau, B. H. (2009). Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited. *Journal of Voice*, 425-438.
- Salodin, C. (2008). *Complete Vocal Technique* (2. ed.). Copenhagen: CVI Publications.
- Salodin, C. (2023, 03 19). *CVT Research*. Retrieved from <https://cvtresearch.com/anatomy-physiology/the-vocal-folds/>
- Sengpiel, E. (13. 03 2023). *Sengpiel Audio*. Von <http://www.sengpielaudio.com/Vokaldreieck.pdf> abgerufen
- Smialek, E. D. (2012). A Spectrographic Analysis of Vocal Techniques in Extreme Metal for Musicological Analysis. *ICMC2012*, 88-93.
- Smialek, E. P. (2012). *Musical aspects of vowel formants in the extreme metal voice*. York, UK: Proc. of the 15th Int. Conference on Digital Audio Effects.
- Sundberg, J. (1977). The Acoustics of the Singing Voice. *Scientific American*, 82-91.
- Sundberg, J. (1987). *The science of the singing voice*. Northern Illinois University Press.
- Sundberg, J. L. (2013). Formant Tuning Strategies in Professional Male Opera Singers. *Journal of Voice*, 27(3), 278-288.
- Winn, M. (22. 12 2020). Abgerufen am 25. 04 2023 von GitHub: https://github.com/ListenLab/make_vowel_space/blob/main/Extract_formants.txt
- Winn, M. (26. 03 2020). *Git Hub*. Abgerufen am 26. 03 2023 von https://github.com/ListenLab/Praat/blob/master/Spectrum_tabulation.txt

-
- Winn, M. (2020, 03 27). *Praat 20 - draw speech spectrum using R*. Retrieved from <https://www.youtube.com/watch?v=clRkjJE252A&list=PLouaSPIGAE9u3yDBcjDphmKHTO9kXE1aq&index=10>
- Winn, M. (2020, 03 25). *Praat 7 - Compare spectra of different speech sounds*. Retrieved from <https://www.youtube.com/watch?v=-tCcQsgZ31E&list=PLouaSPIGAE9u3yDBcjDphmKHTO9kXE1aq&index=7>
- Winn, M. (2021, 02 01). *Make your own vowel chart!* Retrieved from <https://www.youtube.com/watch?v=BGW8J4cG0qY&list=PLouaSPIGAE9u3yDBcjDphmKHTO9kXE1aq&index=5&t=2s>

Anlagen

Anlage 1:

Das vorliegende Praat-Script wurde für die Extraktion von annotierten Zeitbereichen der Vokalformanten verwendet.

```
# Preparation:
# work with a recording that contains the vowels
#   that you want to analyze
#
# Create a formant object for that sound,
# and do it carefully
# (ensure the formant settings are appropriate
#   for the individual talker's voice)
#
# Mark the intervals that you want to analyze
#   in a Textgrid, like this:
#
#   ~~~~~-----
#
#   |  ah  |   |  ih  |   |  eh  |
#
#####
#
# Name of the Textgrid file that you annotated
# (should also be the name of the Formant object)
#   name$ = "02_vowels"
#
# How many timepoints?
#   num_timepoints = 10
#
# Which tier are your annotations?
#   v_tier = 1
#
# Make the table
#   Create Table with column names: "formants", 0, "vowel time_index"
v_time time_abs F1 F2 F3 F4"
#   row_index = 0
#
# Count the intervals
#   select TextGrid 'name$'
#   num_intervals = Get number of intervals: v_tier
#
# Loop through the intervals
#   for interval_index from 1 to num_intervals
#   select TextGrid 'name$'
#   label$ = Get label of interval: v_tier, interval_index
#
#   proceed if the label isn't empty
#   if label$ <> ""
#       t_start = Get start time of interval: 1, interval_index
#       t_end = Get end time of interval: 1, interval_index
#       time_interval = (t_end - t_start)/(num_timepoints-1)
#
#       selectObject: "Formant 'name$'"
#
#   Loop through the timepoints
#       for time_index from 1 to num_timepoints
```

```

time_re_onset = (time_index-1)*time_interval
current_time = t_start + time_re_onset
select Formant 'name$'
f1 = Get value at time: 1, current_time, "hertz",
"Linear"
f2 = Get value at time: 2, current_time, "hertz",
"Linear"
f3 = Get value at time: 3, current_time, "hertz",
"Linear"
f4 = Get value at time: 4, current_time, "hertz",
"Linear"

# Add info to the table
select Table formants
Insert row: row_index + 1
row_index = row_index + 1
Set string value: row_index, "vowel", label$
Set numeric value: row_index, "time_index",
time_index
Set numeric value: row_index, "v_time",
'time_re_onset:3'
Set numeric value: row_index, "time_abs", 'cur-
rent_time:3'
if f1 != undefined
Set numeric value: row_index, "F1", 'f1:0'
else
Set string value: row_index, "F1", "NA"
endif
if f2 != undefined
Set numeric value: row_index, "F2", 'f2:0'
else
Set string value: row_index, "F2", "NA"
endif
if f3 != undefined
Set numeric value: row_index, "F3", 'f3:0'
else
Set string value: row_index, "F3", "NA"
endif
if f4 != undefined
Set numeric value: row_index, "F4", 'f4:0'
else
Set string value: row_index, "F4", "NA"
endif
endif
endfor

# end conditional if label isn't blank
endif

# end loop through the intervals
endfor

```

Anlage 2:

Das folgende Praat-Script wurde für die Tabellierung des Amplitudenspektrums angewendet, um diese später mit der Software R darstellen zu können.

```
# List the values of a spectrum
# in multiple sounds that you select
#
# also with an option to combine all the data
# into one table

clearinfo

# Ask the user to select the sounds they want to use
pause select all sounds to be used for this operation
number_of_selected_sounds = numberOfSelected ("Sound")

# Assign an object number to each of those sounds
for index to number_of_selected_sounds
    sound'index' = selected("Sound",index)
endfor

# Loop through the sounds
for current_sound_index from 1 to number_of_selected_sounds
    select sound'current_sound_index'
        name$ = selected$("Sound")

        # Assign it to a vector of names
        # to call up later
        names$['current_sound_index'] = name$

    ## Do your stuff here
    To Spectrum: "yes"
    Tabulate: "no", "yes", "no", "no", "no", "yes"

    ## Insert the object name into the first column
    selectObject: "Table 'name$'"
    Insert column: 1, "name"
    num_rows = Get number of rows
    for row_index from 1 to num_rows
        Set string value: row_index, "name", name$
    endfor

    ## Make better column names
    Set column label (label): "freq(Hz)", "Frequency"
    Set column label (label): "pow(dB/Hz)", "Intensity"

    select Spectrum 'name$'
    Remove
endfor

if number_of_selected_sounds > 1

# Ask the user if they want to combine all spectra into one single ta-
ble
beginPause ("Combine?")
comment ("Combine all the spectra into a single table?")
combine_into_single_table = endPause ("Cancel", "YES", "NO", 2, 2)
```

```
if combine_into_single_table == 2
# a fake line to initiate the selection of objects
  nocheck select mongoose

# Loop through the sounds
for current_sound_index from 1 to number_of_selected_sounds
  name_string$ = names$['current_sound_index']
  plus Table 'name_string$'

endfor

Append
Rename: "All_spectra"

# Delete the separate objects
# a fake line to initiate the selection of objects
  first_Table$ = names$[1]
  select Table 'first_Table$'
# Loop through the sounds
for current_sound_index from 2 to number_of_selected_sounds
  name_string$ = names$['current_sound_index']
  plus Table 'name_string$'
endfor
Remove

endif

# Note: if you're comfortable with scripting,
# you can automatically save the big table
# with a line like this:
# Save as comma-separated file:
"C:\Users\Matt\Sounds\All_my_sounds.Spectrum"

# end conditional only if there were multiple tables
endif
```

Anlage 3:

Das folgende R-Script wurde für die visuelle Darstellung der Praat-Tabellierung des Amplitudenspektrums angewendet.

```
library(ggplot2)
library(zoo)

# read the Praat spectrogram functions
setwd("D:\\Documents\\Bachelorarbeit\\Aufnahmen\\Praat
Scripts\\spectrum tables")

spectra <- read.csv("combined_creak_spectrum.Table")

#draw spectrum panels
px_spectra_faceted <-
```

```
ggplot(spectra)+
aes(x = Frequency, y = Intensity, color = name)+
geom_line()+
scale_fill_gradient(high="black", low="white", na.value = "white")+
coord_cartesian(ylim=c(-15, 50),
                xlim = c(0, 5000),
                expand = FALSE)+
ylab("Intensity")+
scale_x_continuous("Frequency (Hz)", breaks=seq(0, 5000, 1000))+
theme_bw()+
facet_wrap(~name)
px_spectra_faceted

#draw spectrums to single panel
px_spectra_single_panel <- ggplot(spectra)+
  aes(x = Frequency, y = Intensity, color = name)+
  geom_line()+
  coord_cartesian(xlim = c(0, 5000),
                  ylim = c(0, 50))+
  theme_bw()
px_spectra_single_panel

#smoothing with zoo
spectra$Intensity_smooth <-
  zoo::rollmean(x = spectra$Intensity, k=400, na.pad = TRUE)

#draw spectrum with smoothing
px_spectra_single_panel_sm <- ggplot(spectra)+
  aes(x = Frequency, y = Intensity_smooth, color = name)+
  geom_line()+
  xlab("Frequency (Hz)")+
  ylab("Intensity(dB)")+
  coord_cartesian(xlim = c(0, 5000),
                  ylim = c(-20, 50))+
  theme_bw()
px_spectra_single_panel_sm

#save
ggsave(px_spectra_single_panel_sm,
        file = "combined creak.png",
        height = 3.2, width = 4.8, dpi = 300)
```

Anlage 4:

Diese Tabelle zeigt die Titel, welche im Rahmen der Bachelorarbeit aufgenommen wurden. Die grau markierten Positionen wurden letztendlich für die Klanganalyse ausgewählt.

Take Index	Song	Band	Sänger	Gesangstechnik
2	Agora	Invent Animate	S2	False Cord Scream ("Low")
12	As I Slither	Kataklysm	S3	False Cord Scream ("Low")
24	Blackwater Park	Opeth	S1	False Cord Scream ("Low")
31	Blackwater Park	Opeth	S1	False Cord Scream ("Mid")
8	Atlas	Sentinels	S2	False Cord Scream ("Mid")
17	Three Bridges	Cult of Luna	S3	False Cord Scream ("Mid")
28	Flower of Hate	LXO	S4	False Cord Scream ("Mid")
32	Blackwater Park	Opeth	S1	False Cord Scream ("High")
5	Pain Remains	Lorna Shore	S2	False Cord Scream ("Tunnel Throats" und "High")
6	God knows	Knocked Loose	S2	False Cord Scream ("High")
18	Ruin	Lamb of God	S1	Hyper-Compressed Fry Scream ("High" zu "Low")
19	Ruin	Lamb of God	S1	Hyper-Compressed Fry Scream ("Mid")
20	Can You Feel My Heart	Bring Me The Horizon	S1	Hyper-Compressed Fry Scream ("Pitched")
21	Can You Feel My Heart	Bring Me The Horizon	S1	Hyper-Compressed Fry Scream ("Pitched")
1	Agora	Invent Animate	S2	Loud Fry Scream ("Shout")
4	Linger	Oceans Ate Alaska	S2	Loud Fry Scream ("Shout")
7	God knows	Knocked Loose	S2	Loud Fry Scream ("Shout")
3	Linger	Oceans Ate Alaska	S2	False Cord Scream ("High")
9	Rotations	Reflections	S2	False Cord Scream ("High" und "Mid")
10	Rotations	Reflections	S2	False Cord Scream ("Mid" und "High")
15	Aeons Elapse	Ahab	S3	False Cord Scream ("Low"/"Grunt")

11	Lost in a Loop	Ghost Brigade	S3	False Cord Scream ("Mid" und "Low"/"Grunt")
13	Cathedral Walls	Swallow the Sun	S3	Clean
14	Heart of a Cold White Land	Swallow the Sun	S3	Clean
16	Passing Through	Cult of Luna	S3	Clean
25	Mammoth	King Buffalo	S4	Clean
26	Saturine & Iron Jaw	All them Witches	S4	Clean
27	Not Again	Staind	S4	Clean

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Leipzig, 19.05.2023

Alfred Geier