

# VOCAL CHAINS

Ton Seminar (221300a) WS24/25

eingereicht bei Prof. Oliver Curdt

Felix Schwarz 46007  
fs179@hdm-stuttgart.de

# Inhaltsverzeichnis

1. Einleitung.....	2
2. Probleme von Stimmaufnahmen.....	2
3. Maßnahmen vor der Vocal Chain .....	2
4. Säubern des Stimmsignals.....	3
4.1. Surgical Equalizer .....	3
4.2. De-Esser .....	4
4.3. Kompressor.....	4
5. Gestalten des Stimmsignals.....	7
5.1. Coloring Equalizer .....	7
5.2. Sättigung .....	7
5.3 Sends .....	7
6. Mixing Philosophie und Fazit.....	9

## 1. Einleitung

Die menschliche Stimme ist eines der vielseitigsten und ausdrucksstärksten Instrumente in der Musikproduktion. Um Gesang optimal in einer Produktion zu integrieren, kommen sogenannte Vocal Chains zum Einsatz. Es ist eine Kombination aus verschiedenen Audiotools, die den Klang veredeln, Fehler korrigieren und die gewünschte Klangästhetik formen. Hierbei müssen sowohl technische als auch kreative Entscheidungen getroffen werden. Innerhalb dieser Ausarbeitung werden die verschiedenen Herangehensweisen und wichtigsten Werkzeuge einer Vocal Chain erläutert. Zudem werden verschiedene best Practices und etablierte Tricks zum Bearbeiten von Gesang vorgestellt. Viele der vorgestellten Techniken orientieren sich an der Bearbeitung von Modernen Pop Vocals.

## 2. Probleme von Stimmaufnahmen

Es gibt verschiedene Arten von Problemquellen, die beim Aufnehmen von Vocals vorliegen können. Unterschiedliche Problematiken lassen sich durch tontechnische Eingriffe beheben oder vermindern, einige jedoch nicht. Aufgrund der Aufnahmetechnik entstehen Hintergrundgeräusche sowie Störfrequenzen, da jedes elektronische Gerät rauscht. Durch den Aufnahmeraum entstehen Kammfilter Effekte und eine klangliche Färbung der Stimme. Zudem haben aufgenommene Stimmen oftmals eine sehr hohe Dynamik, sodass sie sich unbearbeitet nur schwer im Gesamtmix durchsetzen können. Kleinere Schwächen in der Stimmenperformance bei Timing und Tuning lassen sich ebenfalls zufriedenstellend durch tontechnische Eingriffe verbessern. Was sich jedoch kaum bis gar nicht durch technische Bearbeitungen verbessern lässt sind emotionale Schwächen in der Performance während der Aufnahme. Somit macht es mehr Sinn eine Sprachaufnahme zu wählen die nur Probleme enthält, die sich tontechnisch verbessern lassen.

## 3. Maßnahmen vor der Vocal Chain

Es gibt verschiedene Schritte und Maßnahmen, welche vor der Bearbeitung durch eine Vocal Chain realisiert werden können. Hierzu zählt das Vocal Comping, sprich das Herausarbeiten der gelungensten Stellen aus jedem Take und das nahtlose Zusammenfügen dieser. Hierzu zählen ebenso manuelles Tuning und Time Stretching, sowie das händische Herausschneiden von Atmergeräuschen, Rauschen und Klickgeräuschen. Zudem können bereits in diesem Prozess einige grobe Dynamikangleichungen durchgeführt werden, sodass für die spätere Effektkette optimalere Threshold Werte gewählt werden können. Zuletzt lassen sich noch harsch klingende Silben und S-Laute herunter regeln um einen maßgeschneiderten De-Esser Effekt zu erhalten.

Viele dieser Schritte können auch innerhalb der Vocal Chain gelöst werden. Beispielsweise kann das Stummschalten von leerem Rauschen auch durch einen Gate Effekt gelöst werden oder das Herunterregeln von S-Lauten durch einen klassischen De-Esser erfolgen. Je nach Ausgangssignal eignen sich manuelle Anpassungen besser oder schlechter. Hierbei sollte man sich für die Arbeitsweise entscheiden, welche zufriedenstellende Ergebnisse mit möglichst geringen langfristigen Arbeitsaufwand realisieren kann.

#### 4. Säubern des Stimmensignals

Wie bereits beschrieben besitzen aufgenommene Stimmensignale verschiedene Stör- und Fehlerquellen. Störfrequenzen, Rauschen und Ähnliches wird durch bestimmte Effekte in der Effektkette verstärkt. Aufgrund dessen ist es sinnvoll im ersten Schritt einer Effektkette das Signal von verschiedenen Artefakten und Störgeräuschen zu säubern, bevor es durch gestalterische Effekte bearbeitet wird. Zudem sollte die Reihenfolge der Effekte vor allem bei Regelverstärkern beachtet werden. So sollte ein Gate immer vor der Eingrenzung der Dynamik mit einem Kompressor verwendet werden, da hier leichter ein optimaler Threshold gewählt werden kann. Zudem sollten Effekte, die zusätzliche Informationen zum Signal addieren meist am Ende der Effektkette verwendet werden. Hierzu zählen beispielsweise Sättigungs Effekte und Hall Effekte. Bei mehreren Stimmen gleichzeitig oder bei zusätzlichen Harmonien einer Stimme erfolgt eine gemeinsame Gruppenbearbeitung am Ende.

##### 4.1. Surgical Equalizer

Ziel eines Surgical Equalizers ist es Störfrequenzen, sowie unästhetische Elemente einer Stimme herauszufiltern. Hierbei werden wünschenswerte und ästhetische Bestandteile der Stimme noch nicht angehoben. Für gewöhnlich verwendet man hierfür einen neutralen digitalen Equalizer. In einem späteren Arbeitsschritt kann man sich somit fokussierter um das Herausarbeiten der wünschenswerten Frequenzbänder kümmern. Klassischerweise wird beim Surgical Equalizer zudem ein Hochpassfilter angewandt um Frequenzbestandteile unter dem Grundton der Stimme herauszufiltern, da diese keine nutzbaren Informationen der Stimme enthalten. Durchschnittlich befindet sich der Grundton von männlichen Stimmen bei ca. 100 Hertz und der Grundton von weiblichen Stimmen bei ca. 200 Hertz. Durch das Anheben verschiedener Frequenzbänder kann nach Resonanz- und Störfrequenzen gesucht werden. Sind diese gefunden können sie durch Notch-Filtern mit verhältnismäßiger Flankensteilheit herausgezogen werden. Bei gewissen Störfrequenzen kann auch das Verwenden eines dynamischen Equalizers sinnvoll sein um den Klang nicht ständig zu beeinflussen.

## 4.2. De-Esser

Oftmals werden S-Laute bei Sprechtaufnahmen als störend oder harsch empfunden. Dieses Problem ist vor allem bei deutschen Sprachaufnahmen verstärkt vorhanden. Die störenden Silben können entweder händisch oder durch einen De-Esser heruntergepegelt werden. Je nach Stimme befinden sich die störenden S-Laute in verschiedenen Frequenzbändern. Oftmals kommt es vor, dass verschiedene S-Laute in verschiedenen Frequenzbereichen dominant sind. Hierbei kann es sinnvoll sein einen breitbandig eingestellten De-Esser zu verwenden und diesen mit einem dynamischen Equalizer zu unterstützen, der verschiedene Frequenzbänder je nach Art des S-Lauts absenkt. Statt eines dynamischen Equalizers kann hier ebenfalls ein zweiter De-Esser mit anderen Frequenzband Einstellungen gewählt werden. Ziel ist es die störenden Silben ohne das Entstehen von Artefakten oder Lispeln herabzusenken. Je nach Genre und Stimme ist eine verschieden starke Absenkung erwünscht. Bei moderner Pop Musik wird oftmals eine starke Absenkung angewandt, da die Stimme innerhalb dieses Genres oft durch starke Anhebung innerhalb der hohen Mitten und Höhen präsenter gemacht wird. Wäre hierbei kein starkes De-Essing vorhanden würde die Stimme schnell zu scharf klingen.

## 4.3. Kompressor

Eines der wichtigsten Werkzeuge der Tontechnik zur Signalverarbeitung ist der Kompressor. Innerhalb der Geschichte der modernen Tontechnik wurden verschiedene Arten von Kompressoren entwickelt. Hierbei haben sich viele Herangehensweisen und Workflows mit verschiedenen Arten von Kompressoren etabliert. Diese Herangehensweisen können heutzutage mithilfe von Emulationen analoger Hardware, sowie dem sinnngemäßen Imitieren mithilfe von digitalen Kompressoren angewandt werden. Die verschiedenen Arten von Kompressoren die hierbei vorgestellt werden sind die Opto-, FET-, VCA- und Digital-Kompressoren.

Die ersten Kompressoren die für den tontechnischen Einsatz entwickelt wurden sind die Röhren Kompressoren. Sie wurden vor allem im Rundfunk eingesetzt. Ihre Bautechnik funktioniert dadurch, dass eine anliegende Bias Spannung den Verstärkungsfaktor einer elektrotechnischen Röhre beeinflusst. Hierdurch wird die Kompression realisiert. Da eine Röhre ein hochgradig nichtlineares Bauteil darstellt entstehen hierdurch Verzerrungen, die zu einer starken Färbung des Klanges führen. Zudem ist die angewandte Ratio abhängig von der Lautstärke des Eingangssignals. Dies ist vergleichbar mit einem digitalen Kompressor mit hoher Soft Knee Einstellung. Dies führt zu einem charakteristischen Klang von Röhren Kompressoren der sich heutzutage vor allem zum stilistischen Gestalten von Stimmen eignet. Röhren Kompressoren haben vergleichsweise langsame Regelzeiten und ein intransparentes Klangbild. Hierdurch eignen sie sich nur schlecht zum Abfangen von Transienten und schnellen Pegelspitzen einer Stimme. Ein bekanntes Beispiel für einen Röhren Kompressor ist der Fairchild 670.

Die Kompressoren, welche als nächstes entwickelt wurden sind die so genannten Opto Kompressoren. Das Eingangssignal beeinflusst in ihrer Schaltungstechnik dem Helligkeitsgrad einer verbauten Glühbirne oder Leuchtdiode. Je lauter das Eingangssignal ist, desto heller leuchtet die Glühbirne/Leuchtdiode. Neben der Glühbirne/Leuchtdiode befindet sich ein Photowiderstand, der seinen Widerstandswert abhängig der Helligkeit der Glühbirne/Leuchtdiode verändert. Je mehr Licht abgestrahlt wird, desto höher ist der Widerstandswert der das Ausgangssignal beeinflusst. Hierdurch wird die Kompression eines Opto Kompressors realisiert. Das Klangbild eines Opto Kompressors wird durch seine mittelschnelle Attacke Zeit und vor allem seine sehr langsame Release Zeit charakterisiert. Hierdurch ergibt sich ein vergleichsweise transparentes Klangbild. Verschiedene Opto Kompressoren haben zudem oftmals viele nichtlineare Bauteile verbaut, weshalb diese Kompressoren ebenfalls eine Verzerrung und Klangfärbung hervorrufen. Auch wenn Opto Kompressoren ebenfalls Schwierigkeiten haben schnelle Pegelspitzen abzufangen, eignen sie sich hervorragend einer Stimme durch Färbung und langsamer Dynamikangleichung eine hohe Präsenz und Durchsetzungsfähigkeit im Gesamtmix zu verleihen. Ein bekanntes Beispiel für einen Opto Kompressor ist der LA-2A.

Nach der Entwicklung der Opto Kompressoren folgte die Entwicklung der FET-Kompressoren. FET steht hierbei für Field Effekt Transistor, sprich Feldeffekttransistor, welcher ebenfalls das zentrale elektrotechnische Bauteil darstellt. Der Feldtransistor ändert seinen Widerstandswert abhängig vom Eingangssignal und wendet diesen auf das Ausgangssignal an. Hierdurch wird eine Kompression mit sehr schnellen Regelzeiten realisiert. FET-Kompressoren haben hierdurch ein stark intransparentes Klangbild, welches bei schnellen Regelzeiten zu starken Verzerrungen und Färbungen führt. Eine starke Kompression ist deutlich hörbar. FET Kompressoren eignen sich Hervorragend um Pegelspitzen abzufangen, Parallelkompression zu realisieren oder Effekte umzusetzen bei denen eine stark hörbare Kompression erwünscht ist. Ein bekanntes Beispiel für einen FET Kompressor ist der 1176.

Die neueste relevante Art analoger Kompressoren sind die VCA Kompressoren. VCA steht hierbei für Voltage Controlled Amplifier, was eine Schaltung darstellt die mithilfe von mehreren Transistoren realisiert wird. VCA Kompressoren zeichnen sich durch ihre hohe Flexibilität aus. Sie können sowohl schnelle als auch langsame Regelzeiten mit hoher Transparenz realisieren. Im Vergleich zu anderen analogen Kompressoren haben VCA Kompressoren nur einen geringen Verzerrungsgrad und damit nur eine geringe Klangfärbung. VCA Kompressoren können somit sehr vielseitig eingesetzt werden. Ihnen fehlt es jedoch an klanglichen Charakteristiken, welche die anderen Kompressoren ausmacht. Ein bekanntes Beispiel für einen VCA Kompressor ist der Distressor.

Mit der Entwicklung von digitaler Audiotechnik entstand eine neue Art von digitalen Kompressoren. Hierbei wird die Kompression mithilfe digitaler Bearbeitung durch eine

Programmiersprache realisiert. Hierdurch ist eine maximale Flexibilität von Ratio und Release Zeiten möglich. Digitale Kompressoren besitzen in der Regel keine der Kompression nicht inhärente Art der Verzerrung und eine hohe Transparenz. Somit fügen sie dem Signal keine zusätzliche Klangfärbung hinzu.

Beim Verwenden von Kompressoren in einer Vocal Chain ist es wichtig zu beachten welche Art von Kompressor verwendet wird. Ziel ist es je nach Stimme und Genre die Stärken eines Kompressors zu nutzen um ein gewünschtes Ergebnis zu erhalten. Die Entwicklung von digitalen Kompressoren und die Möglichkeit mit hoher Flexibilität und geringer Klangfärbung Kompression zu realisieren ermöglicht älteren Arten der Kompression ein gestalterisches Mittel zu sein um bestimmte Klangästhetiken zu realisieren. Verschiedene Kompressoren können miteinander kombiniert werden um ihre Schwächen zu kaschieren oder ihr einzigartiges Klangbild zu entfalten.

Will man lediglich die Dynamik einer Stimme grob kontrollieren, genügt es in der Regel einen Kompressor zu verwenden, der mit einer Ratio von ca. 3:1 bis 4:1 das Signal um ca. 3dB reduziert. Hierbei sollte die Attack Zeit möglichst kurz gewählt werden ohne, dass es zu Verzerrungen kommt, damit der Kompressor schnelle Pegelspitzen abfangen kann um Clippen im Gesamtmix zu verhindern. Für viele Stimmensignale reicht diese Art der Kompression problemlos aus, dennoch ist hiermit nur eine begrenzte Dynamikreduktion ohne Artefakte möglich.

Für eine starke Durchsetzbarkeit und Dynamikreduktion werden für gewöhnlich zwei oder mehrere Kompressoren seriell hintereinander verwendet. Der erste Kompressor besitzt die Aufgabe schnelle Pegelspitzen zu reduzieren ohne das restliche Signal stark zu beeinflussen. Hierfür eignen sich FET-, VCA- oder digitale Kompressoren mit schnellen Regelzeiten und hoher Ratio. Der Threshold sollte hierbei vergleichsweise hoch gewählt werden. Die Rolle des zweiten Kompressors ist es das von Pegelspitzen bereinigte Signal ständig leicht zu komprimieren um ein sich stark durchsetzendes Signal zu erhalten. Hierfür eignen sich Röhren-, Opto- oder digitale Kompressoren mit geringer Ration und langsamen Regelzeiten. Hierbei kann ein explizit färbender Kompressor helfen die stimme klanglich zu gestalten. Diese Herangehensweise von serieller Kompression ist vor allem bei Pop und Hip-Hop Produktionen beliebt um eine Stimme stark im Vordergrund eines Mixes platzieren zu können.

Wenn eine explizit sanfte Art der Kompression gewünscht ist gibt es eine Herangehensweise, die sich in der Nutzung des Kompressors von SSL-Mischpulten etabliert hat. Diese Herangehensweise ist vor allem bei Indie, Jazz und Soul Musik beliebt. Hierbei setzt man den Threshold des Kompressors sehr niedrig, sodass der Kompressor das Signal ständig komprimiert. Jedoch wird eine sehr geringe Ratio von ca. 1.2:1 bis 1.5:1 verwendet, sodass nur eine geringfügige Kompression umgesetzt wird. Hierdurch wird eine hohe Transparenz erzielt.

## 5. Gestalten des Stimmsignals

Nachdem das Stimmensignal von Störfaktoren bereinigt und für die weitere Bearbeitung aufgewertet wurde, kann das Signal gestalterisch verfeinert werden. Ziel hiervon ist es die ästhetischen Merkmale einer Stimme herauszuarbeiten und passend für eine gewisse Klangästhetik oder ein gewisses Genre anzupassen.

### 5.1. Coloring Equalizer

Ziel eines Coloring Equalizers ist es durch breitbandige Anpassungen im Frequenzspektrum die wünschenswerten Frequenzbänder eines Stimmensignals herauszuarbeiten. Anders, als bei einem Surgical Equalizer kann hier auch auf analoge Equalizer zurückgegriffen werden. Für gewöhnlich versucht man mit einem Glockenfilter die Boominess (ca. 100-300 Hz) und Boxyness (ca. 400-500 Hz) zu kontrollieren. Eine nasal klingende Stimme kann durch Anpassungen zwischen 1 – 4 kHz Klarheit verschafft werden. Bei moderner Pop Musik ist in diesem Schritt ein starker Boost in den Höhen und Hohen Mitten üblich.

### 5.2. Sättigung

Viele analoge Effektgeräte und Mischpulte haben sich dadurch ausgezeichnet, dass sie dem Signal aufgrund nichtlinearer Bauteile durch Verzerrung und Sättigung Obertöne hinzugefügt haben, die als besonders ästhetisch wahrgenommen wurden. Diese Sättigung auf Stimmensignalen hat sich im Laufe der Musikgeschichte zu einer Hörerwartung entwickelt. In der Arbeitsweise mit digitaler Signalverarbeitung fehlen diese Obertöne größtenteils. Aufgrund dessen ist es üblich einem digital bearbeiteten Stimmensignal künstlich Verzerrung hinzuzufügen. Dies kann mit Emulationen von analogen Preamps, Channel Strips oder anderweitigen Distortion Effektgeräten realisiert werden. Ob und wie stark Sättigung auf das Stimmensignal hinzugefügt wird hängt stark vom zu realisierenden Genre ab. Vorteile von Sättigung auf Stimmensignalen stellt eine Verdichtung des Frequenzspektrums, sowie eine Erhöhung der wahrgenommenen Lautstärke dar. Dies macht die Platzierung einer Stimme im Gesamtmix in der Regel einfacher.

### 5.3 Sends

Am Ende einer Vocal Chain befinden sich die Effekte, welche auf den Send Kanälen liegen. Für gewöhnlich sind dies verschiedene Ear Candy Effekte, Parallelbearbeitungen, Delays und Hall.

Vor allem Hall wird bei fast keiner modernen Vocal Chain ausgelassen, da die meisten Aufnahmen in trockenen, schalltoten Räumen aufgenommen werden und ohne künstlichen Hall unnatürlich und unästhetisch klingen. Es gibt verschiedene Arten von künstlichem Hall, die sich für verschiedene Stilrichtungen eignen.

Am universell einsetzbaren ist der Room-Reverb, welcher das Nachhallverhalten eines generischen Raumes simuliert. Normalerweise ist hierbei die Größe des Raumes, sowie die Nachhallzeit einstellbar. Diese Art des Halls liefert einen natürlichen und vertrauten Klang welcher somit für viele verschiedene Genre und Stimmen einsetzbar ist.

Ähnlich zum Room Reverb ist der Hall-Reverb die Simulation eines Raumes. Statt eines generisch mittelgroßen Raumes wird hier jedoch ein großer Konzertsaal simuliert. Wichtige Charakteristiken eines Konzertsaals für das Nachhallverhalten sind hierbei vor allem die Größe des Raumes, hohe Decken und viele reflektierende Oberflächen. Hierdurch entstehen lange Nachhallzeiten, sowie ein angedicktes Klangbild, welches bei starkem Einsetzen zu Schwierigkeiten im Gesamtmix führen kann. Der Hall-Reverb eignet sich vor allem für orchestrale Performances und kurzes Ear Candy.

Eines der ersten künstlichen Methoden Hall zu simulieren ist der Spring-Reverb. Die elektrischen Stimmensignale werden hierbei durch eine Metallfeder geschickt. Hierdurch entsteht ein besonderer Nachhalleffekt mit speziellem Tremolo in der Hallfahne. Einen Spring-Reverb macht eine künstliche und höhenbetonte Klangästhetik aus. Für eine durchgängige Verwendung ist ein Spring-Reverb oftmals zu aufdringlich. Dennoch eignet sich ein Spring-Reverb für bestimmte Ear Candy Effekt und Reverb Throws.

Eine weitere frühe Methode Hall zu simulieren ist der Plate-Reverb. Hierbei wurden die elektrischen Signale durch eine große Metallplatte geschickt und durch Pickups an den Rändern der Platte wieder aufgenommen. Hierdurch entsteht ein künstlich klingender, heller und nostalgischer Klang welcher ausschlaggebend für bestimmte Genres war. Um eine gewisse nostalgische Klangästhetik zu erzeugen eignet sich der Plate-Reverb hervorragend auf Stimmen.

Mit einem Convolution-Reverb lassen sich mittels Impulsantwort verschiedener Räume und Effektgeräte nachbilden. Hiermit ist es ebenfalls möglich eigene, surreale Reverb Arten zu erzeugen.

Viele Hallarten erzeugen hohe und tiefe Frequenzen, die der Durchsetzbarkeit und Klarheit der Stimme im Gesamtmix schaden können. Mit einem Hochpass-, Tiefpass- oder Bandpassfilter auf dem Hall Send Kanal kann die Klarheit der Stimme verbessert werden. Werden lange Nachhallzeiten und große Räume verwendet kann schnell ein matschiges Klangbild entstehen. Um dem entgegenzuwirken kann man Voice Over Kompression auf dem Hall Send Kanal verwenden, der das Hallsignal während der Gesangspassagen herunterpegelt, sodass hauptsächlich die Hallfahne bestehen bleibt.

## 6. Mixing Philosophie und Fazit

Eine gut gestaltete Vocal Chain kann eine Aufnahme entscheidend verbessern. Durch die Kombination aus technischen Korrekturen und kreativen Effekten wird die Stimme optimal in den Mix eingebunden und klingt professionell und ausdrucksstark. Eine zentrale Erkenntnis ist, dass je nach Stimme und Genre unterschiedliche Entscheidungen getroffen werden müssen. Man sollte jede Vocal Chain von neu anfangen und die Werkzeuge die zur Verfügung stehen individuell einsetzen. Bei vielen Vocal Chains ist zudem weniger mehr. Je weniger Schritte eine Vocal Chain benötigt, desto schneller lässt sie sich realisieren und desto leichter lassen sich Anpassungen vornehmen. Wichtig ist auch welche Werkzeuge zur Realisierung einer Vocal Chain verwendet werden. Prinzipiell lässt sich viel mit einfachen Tools umsetzen, dennoch kommt man mit maßgeschneiderten Tools oftmals schneller an ein zufriedenstellendes Ergebnis. Abschließend gibt es die Erkenntnis, dass keine perfekte oder universell einsetzbare Vocal Chain existiert und je nach zu erzielendem Ergebnis und subjektiven Geschmack aufgebaut werden muss.