



Shure SM7

## Audio für Youtube

### Folge 1

# Sprachaufnahme Off-Stimme

In den letzten Jahren ist die Qualität von Youtube-Produktionen spürbar gestiegen. Viele Kanäle produzieren mit aufwendiger Studiobeleuchtung in 4K, und auch die Audioqualität erreicht bei einigen Vertretern ein hohes Niveau. In dieser Workshop-Serie zeigen wir Ihnen, wie Sie auch ohne ein teures Tonstudio, professionelles Audio für Ihre Youtube-Videos produzieren.

von Björn Eichelbaum

**D**as Imagevideo mit Killer-Animationen ist fertig geschnitten, die Hintergrundmusik pumpt futuristische Stimmung aus den Lautsprechern, die perfekt mit der Schnittfrequenz harmoniert. Doch dann setzt die Sprecherstimme aus dem Off ein: Sie klingt, als ob sie in einer öffentlichen Toilette aufgenommen wurde. Mal wird sie von der Musik verdeckt, mal plärrt sie zu laut. Hinzu kommt noch, dass die meisten Youtube-Videos insgesamt wesentlich lauter und druckvoller klingen.

Was ist da los? Wie kriegt man es hin, dass die eigenen Sprachaufnahmen nach Jakobs Krönung mit Verwöhnaroma klingen? Die Antwort liegt nicht in sündhaft teuren Mikrofonen oder speziellen Audio-Plug-ins. Die werden Sie zwar für die finale Bearbeitung benötigen, entscheidend sind allerdings das richtige Mikrofon in korrekter Aufstellung in einem gut klingendem Raum und nicht zuletzt die Fähigkeiten des Sprechers.

### Akustik

Der Klang einer Sprachaufnahme hängt zu großen Teilen davon ab, wie der Raum, in dem der Sprecher sich befindet, klingt und wie weit er vom Mikrofon entfernt ist. Die extrem nah und dicht klingenden Off-Stimmen werden meist direkt auf die Membran des Mikrofons gesprochen, um so wenig Raumanteil wie möglich mit aufzuzeichnen. Hierbei kommt dem Sprecher der sogenannte Nahbesprechungseffekt zu Hilfe, der bei geringem Sprechabstand zur Membran den Bassbereich einer Stimme anhebt und sie somit voluminöser und gleichzeitig intimer erscheinen lässt.

### Es mikrofoniert!

Ein weiterer Faktor ist die Richtcharakteristik des Mikrofons. Ein Mikrofon mit der Richtcharakteristik Kugel nimmt Schall aus allen Richtungen gleich laut auf. Mikrofone mit der Richtcharakteristik Niere hingegen zeichnen verstärkt in einem schmalen, nierenförmigen Bereich vor dem Mikrofon auf und blenden Schall aus allen übrigen Richtungen weitgehend aus. Für Off-Stimmen sind also Mikrofone mit Niere oder sogar Super- oder Hyperniere besonders gut geeignet.

Doch auch mit Hyperniere und extrem geringem Sprechabstand wird man in einem Raum mit Parkett und ungedämmten Wänden keine vernünftige Off-Stimme hinbekommen. Das liegt an den in jedem akustisch unbehandelten Raum auftretenden Reflexionen, Flatterechos und Raumresonanzen, auch Raummoden genannt. Flatterechos treten besonders an gegenüberliegenden glatten Wänden auf – wie ein Ping-Pong-Ball werden die Schallwellen von einer glatten Wandfläche zur anderen reflektiert. Klatschen Sie einfach in der Mitte des Raumes laut in die



Akustikenelemente nehmen den Schwung aus der Sache – das Echo bleibt aus.

Hände und achten Sie auf die auftretenden Reflexionen – wenn Sie einen metallischen Nachklang mit viel Reflexion hören, sollten Sie die direkt gegenüberliegenden Wandflächen mit Absorbieren und Diffusoren bedecken – das gilt auch für Decke und Boden. Ideal wäre ein schalltoter Raum ohne Resonanzen oder zumindest ein akustisch optimierter Raum für Sprachaufnahmen, wie es ihn in professionellen Tonstudios gibt. Doch so ein Raum ist teuer. Aber es gibt bezahlbare Alternativen. Zum einen können Sie mit relativ geringem Kosten- und Materialaufwand einen kleineren Raum oder eine Abstellkammer mit Akustikelementen auskleiden – meist reicht schon eine stille Ecke, die mit einem dicken, schweren Vorhang vom Rest des Raums abgetrennt werden kann.

Oder Sie kaufen sich für etwas über 100 Euro einen sogenannten Mic Shield, der den Raumschall mit einer halbkreisförmigen Schale aus Akustikschaum dämmt. Für ein paar Hundert Euro mehr bekommen Sie fast komplett geschlossene Systeme, in die der Sprecher seinen Kopf hineinsteckt und die über Klappen verschlossen werden. Diese Systeme bieten die Vorteile einer Sprecherkabine, sind mit Preisen um die 800 bis 900 Euro allerdings wesentlich günstiger und benötigen wenig Platz.

## Mikrofontypen

Für die Aufzeichnung von Off-Stimmen eignen sich Kondensatormikrofone mit Großmembran am besten. Sie können auch mit

dynamischen Mikrofonen wie etwa dem Shure SM7, das im Broadcast- und Studiobereich weit verbreitet ist, gute Ergebnisse erzielen. Ein gutes Kondensatormikrofon bietet zwar etwas mehr Auflösung in Dynamik und Frequenzbereich, benötigt dafür aber eine externe Stromversorgung über Phantomspeisung oder Batterie und ist mechanisch wesentlich empfindlicher und teurer als ein dynamisches Mikrofon.

Überlegen Sie vor dem Kauf, wo und wie Sie das Mikrofon aufstellen wollen. Wenn Sie ein Tischstativ direkt vor dem Monitor aufstellen, sollten Sie bedenken, dass ein Mikrofon ohne elastische Aufhängung – einer sogenannten Spinne – anfällig für Körperschall ist. Sie müssen während der Aufnahme nur einmal kurz an den Tisch stoßen und es rumpelt heftig.

Besser Sie hängen das Mikrofon mit einem großen Stativ über Kopf vor dem Sprecher in etwa 10 bis 15 Zentimetern Abstand auf. Achten Sie darauf, dass der Mund nicht direkt auf die Membran zeigt, sondern das Mikrofon leicht abgewinkelt etwas über der Oberlippe platziert wird. So vermeiden Sie, dass die stark gerichteten Lippen- und Schmatzgeräusche überlaut in Erscheinung treten.

Ein sogenannter Plop-Schutz zur Unterdrückung von Plosivlauten ist bei der Aufnahme von Off-Stimmen aufgrund des geringen Sprechabstands unumgänglich. Sie können sich entweder aus einem Drahtbügel und einer Damenstrumpfhose selber einen Plop-Schutz basteln oder ab etwa 15 Euro einen vorgefertigten im Musikgeschäft kaufen.

## Mikrofonvorverstärker

Leider können Sie ein Kondensatormikrofon nicht so einfach an eine handelsübliche Soundkarte anschließen. Denn zum einen passt das dicke, dreiadrige XLR-Mikrofonkabel nicht in die Miniklinkenbuchse und zum anderem benötigt ein Kondensatormikrofon bis zu 48 Volt Versorgungsspannung, auch Phantomspeisung genannt, wie sie etwa spezielle Soundkarten mit integrierten





**Samson G-Track – USB-Komplettlösung für Einsteiger**

Das USB-Mikrofon Samson G-Track vereint für knapp über 100 Euro alles, was man für eine gute Sprachaufnahme braucht, in einem Gehäuse: Kondensatormikrofon, Kabel, Mikrofonvorverstärker, Soundkarte mit latenzfreiem Monitoring, Kopfhörerverstärker und Tischmikrofonständer. Die Handhabung ist denkbar einfach: Mikrofon via USB an den Computer anschließen, in der DAW- oder Videoschnittsoftware den Treiber wählen, am Mikrofon Aufnahme und Kopfhörerpegel einstellen und aufnehmen.

Neben dem Mikrofonsignal kann über eine Miniklinkenbuchse auch ein Line-Signal eingespeist werden. Eine mehrfarbige LED dient als Übersteuerungsanzeige. Das G-Track bietet latenzfreies Monitoring mit separatem Lautstärkereglern und wird mit einer Vielzahl praktischer Adapterkabel und einem stabilen Tischmikrofonständer ausgeliefert – eine elastische Aufhängung ist optional erhältlich. Klanglich eignet sich das G-Track aufgrund seines angenehm dosierbaren Nahbesprechungseffekts gut für Sprachaufnahmen. Lediglich das Grundrauschen ist etwas hoch, was bei beherrzter Nachbearbeitung mit Kompressor problematisch werden kann.

Hören Sie das Samsong G-Track im direkten Vergleich zu anderen Mikrofonen im Digital Production Youtube-Video zu dieser Folge!

<b>Preis</b>	ca. 110 Euro
<b>Vertrieb</b>	Samson
<b>Internet</b>	<a href="http://www.samsontech.com">www.samsontech.com</a>



Mikrofonvorverstärkern bieten. Eine interessante Alternative bieten Komplettlösungen wie etwa das im Kasten erwähnte Samson G-Track, das via USB an den Rechner angeschlossen wird und Mikrofon, Mikrofonvorverstärker, Soundkarte und Monitoring in einem Gerät vereint. Diese praktische und preiswerte Lösung ist besonders für Einsteiger interessant.

Wesentlich flexibler sind Sie allerdings mit einer Kombination aus Soundkarte mit Mikrofonvorverstärker, die darüber hinaus auch latenzfreies Monitoring mit speziellen Treibern sowie regelbare Kopfhörerausgänge und meist auch attraktive Software-Bundles beinhalten. So bieten etwa die Scarlett-2nd-Generation-Karten von Focusrite ein breit gestaffeltes Angebot von einem bis zu acht Mikrofonkanälen in guter Qualität zu vernünftigen Preisen.

**Kopfhörer & Monitoring**

Für einen Sprecher ist ein gutes Monitoring über Kopfhörer enorm wichtig, weil er fei-

ne sprachliche Nuancen und Betonungen nur dann exakt steuern kann, wenn er sich selbst gut hört.

Auch die durch den geringen Sprechabstand auftretenden Störgeräusche, wie Schmatzer oder Lippengeräusche können durch Ändern der Sprechrichtung und des Abstands fein moduliert werden, wenn der Sprecher hört, was er tut.

Dazu benötigt er einen angenehm sitzenden, geschlossenen Kopfhörer mit guter Dämmung des austretenden Schalls. Denn Sie möchten ja nur die Sprache aufnehmen und nicht, was aus dem Kopfhörer plärrt. Bewährte Klassiker in diesem Bereich sind

der Beyerdynamik DT 100/770 oder der Sony MDR-7506.

Achten Sie bei der Anschaffung der Soundkarte oder einer Komplettlösung auf jeden Fall auf latenzfreies Monitoring. Ohne latenzfreies Monitoring wandelt die Soundkarte das analoge Sprachsignal in ein digitales Signal, das durch diverse Buffer bis in die Software gelangt, dort auf den Kopfhörerausgang geroutet wird und wieder durch Buffer zum Digital-Analog-Wandler der Soundkarte und von dort aus zum Kopfhörerausgang gelangt. Je nach Rechnerhardware und Treiber kann dieser Vorgang bis zu 200 Millisekunden dauern.

Es gibt wenig Dinge, die so irritierend sind, wie sich selbst mit deutlich hörbarer Verzögerung sprechen zu hören – Latenz auf dem Kopfhörerausgang darf nicht sein. Soundkarten mit latenzfreiem Monitoring bieten die Möglichkeit, das Mikrofonsignal direkt und ohne Verzögerung auf den Kopfhörerausgang zu mischen.

Eventuell in der Audiosoftware vollzogene Bearbeitungen via Equalizer oder Kompressor sind dann allerdings nicht ohne Latenz hörbar.





## Vorbereitung und Aufnahme

Eine gute Vorbereitung wird Ihnen während der Sprachaufnahme und nachher bei Schnitt und Bearbeitung viel Zeit sparen. Egal ob Sie selbst oder jemand anderes den Text einspricht, machen Sie es ihm oder sich selbst so leicht wie möglich. Drücken Sie den zu sprechenden Text groß genug aus und fügen Sie nach jedem Satz einen Absatz ein, damit das Satzende deutlich erkennbar ist. Besonders für Anfänger ist es eine gute Idee, den Text vor der ersten Aufnahme zwei- oder dreimal laut durchzulesen und sich bei problematischen Stellen Notizen oder Zeichen zu Betonungen und Pausen zu machen.

Wenn Sie selbst sprechen, versuchen Sie die Off-Stimme mit wenig Dynamik einzusprechen. Lieber etwas leiser, dafür aber kontrollierter und näher am Mikrofon einzusprechen, das vermittelt Intimität und wirkt souveräner als ein Marktschreier. Keine Sorge, durch die später folgende Bearbeitung mit Equalizer und Kompressor wird die Sprache durchsetzungsfähiger klingen. Achten Sie auf ein gleichmäßiges Sprechtempo; besonders bei den ersten Aufnahmen neigt man dazu, immer schneller zu werden.

Machen Sie nach jedem Satz eine Pause und atmen Sie ein und aus bevor Sie mit dem nächsten Satz beginnen. Bei Fehlern empfiehlt es sich, den kompletten Satz, unter Umständen sogar den ganzen Absatz neu einzusprechen. Zwar können Sie nachher im Schnitt die Fehler auch mitten im Satz ganz einfach herausschneiden, ob dann aber die Betonungen und der Sprechrhythmus noch stimmig sind, ist eine andere Frage.

Bereits während des Übens sollten Sie das Pegelmeter des Mikrofoneingangs beobachten und den Pegel anpassen. Achten

Sie darauf, dass der Pegel nicht zu niedrig, aber bei lauten Stellen auch nicht zu hoch ist, sonst gibt es Übersteuerungen und die Aufnahme ist futsch. Wenn Sie sich zwischen  $-25$  dB und  $-6$  dB bewegen, machen Sie nichts falsch. Sollten Sie zum Bild aufnehmen, hilft es, zusätzlich zum Bildstart ein akustisches Signal, etwa einen kurzen Piepston, auf den Kopfhörer zu geben. So kann sich der Sprecher ganz auf den Text konzentrieren und muss nicht ständig auf das Bild achten.

Bei längeren Texten sollten Sie dem Sprecher ein Glas Wasser oder eine Apfelschorle bereitstellen. Die leicht säuerliche Apfelschorle sorgt für einen gewissen Speichelfluss und hält so den Mund und die Lippen angenehm feucht. Kaffee hingegen trocknet den Mund eher aus und sollte daher vor oder nach der Aufnahme getrunken werden.

Überprüfen Sie nach der Aufnahme sorgfältig, ob Sie auch jeden Satz komplett und mit der richtigen Betonung vorliegen haben. Besonders bei vielen Korrekturen während der Aufnahme geht einem gerne mal ein Satzteil verloren. Lieber gleich eine zweite Version einsprechen, als nachher beim Schnitt festzustellen, dass man wegen eines Satzes noch mal alles aufbauen muss oder sogar den Sprecher erneut buchen muss.

## Noise Gate/Schnitt

Je nachdem wie laut die Störgeräusche im Hintergrund zu hören sind, kann es notwendig sein, die einzelnen Sprachtakes hart zu schneiden. Das bedeutet, dass Sie die Wellenform direkt am Wortende und Anfang ohne Vorlauf schneiden. Besonders wenn Sie das Sprachsignal stärker mit Equalizer und Kompressor nachbearbeiten, werden Hintergrundgeräusche,

die Sie während der Aufnahme kaum wahrgenommen haben, unangenehm in den Vordergrund rücken. Alternativ können Sie die Hintergrundgeräusche auch mithilfe eines Noise Gates automatisch ausblenden. Stellen Sie dazu einfach den Parameter „Threshold“ so ein, dass das Gate in Pausen schließt und bei Sprache öffnet. Das funktioniert bei weniger stark komprimierten Sprachsignalen meist ganz gut. Sobald Sie stärker komprimieren werden Sie um das Zuschneiden von Hand nicht mehr herumkommen.

## Kompressor

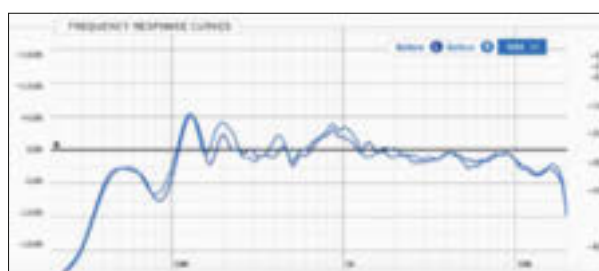
Mit einem Kompressor schränken Sie die Dynamik des Sprachsignals ein – vereinfacht ausgedrückt werden die lauten Stellen leiser und die leisen Stellen dadurch im Verhältnis lauter. Dadurch, dass die leisen Anteile nun besser zu hören sind, entsteht für unseren Gehörsinn der Eindruck, die Stimme sei näher am Ohr. Schauen Sie nach, ob der Audio-Kompressor Ihrer DAW- oder Videoschnitt-Software über ein Preset „Speech“ oder „Vocal“ verfügt und versuchen Sie es mit dieser Einstellung.

Sonst probieren Sie folgende Einstellung aus: Attack 8-15 ms, Ratio 1:8, Release Auto/300-500 ms. Ziehen Sie bei dieser Einstellung den Threshold-Regler nun so lange herunter, bis das Meter „Gain Reduction“ eine Pegelreduktion von  $-6$  dB bis  $-20$  dB anzeigt. Den eventuell auftretenden Pegelverlust kompensieren Sie über „Output Gain“ oder, falls Ihr Kompressor diese Funktion bietet, über „Auto Gain“.

## Equalizer

Mithilfe des Equalizers werden Sie einerseits unerwünschte und störende Frequenzen absenken und andererseits die Frequenzen anheben, die das Sprachsignal noch intimer und wärmer klingen lassen. Trotz guter Dämmung verfügen die meisten Aufnahme Räume über die eingangs erwähnten Resonanzen/Raummoden, die auch auf der Aufnahme zu hören sein werden. Durch die Kompression treten diese nun noch stärker in den Vordergrund, weshalb Sie diese mit dem EQ absenken müssen.

Heben Sie bei einer Männerstimme zunächst bei 100 Hz etwa zwei bis drei dB





breitbandig an. Als nächstes wählen Sie ein parametrisches oder auch Peak- oder Bell-EQ genanntes Band, wählen als Frequenz 200 Hz, einen niedrigen Q-Faktor – also ein schmales Band – und heben Sie dieses um etwa 12 dB an. Wenn Sie nun die Frequenz des EQ langsam zwischen 100 und etwa 250 Hz hin- und herbewegen, werden Sie bei einer bestimmten Frequenz ein Wummern bemerken. Hier liegt die störende

Resonanzfrequenz. Senken Sie diese um 6 bis 12 dB ab und der Bassbereich sollte voluminös, aber trotzdem nicht dumpf klingen.

Um den Eindruck der Intimität zu erhöhen und der Stimme etwas Glanz zu verleihen, heben Sie mit einem High Shelf EQ das Höhenband ab 6 bis 9 KHz um zwei bis drei dB an. Je nach Charakter und Geschlecht der Stimme können Sie mit einer breitbandigen Anhebung zwischen 3,5 und 5 KHz um etwa 2 bis 4 dB die Sprachverständlichkeit erhöhen. Jede Stimme und Aufnahmesituation ist einzigartig, deshalb sollten Sie die genannten Werte nur als Orientierung verstehen. Wenn Ihre Aufnahme mit extremen Einstellungen gut klingt, spielt das keine Rolle, entscheidend ist, wie es klingt.

### De-Esser & Maximizer

Der Kompressor kann neben den Resonanzen im Bassbereich auch bestimmte „S“- und Zischlaute unangenehm in den Vordergrund rücken. Durch das Hinzufügen von reichlich Glanz im Höhenbereich verschlimmert sich dieser Effekt meist noch.

Abhilfe schafft hier ein Plug-in mit dem bezeichnenden Namen De-Esser. Ähnlich

wie beim Kompressor steuern Sie die Pegelreduktion auch beim De-Esser mit dem Threshold-Regler. Versuchen Sie, mithilfe des Bandwidth-Reglers, die „S“-Frequenz so genau wie möglich einzugrenzen, um möglichst wenig Glanz zu verlieren.

Zum Abschluss können Sie das Sprachsignal zusätzlich mit einem sogenannten Maximizer noch lauter und durchsetzungsfähiger gestalten. Je nach Plug-in, müssen Sie dazu meist nur einen Threshold-Regler herunterziehen oder den Parameter „Maximize“ hoch regeln. In der nächsten Folge beschäftigen wir uns mit Sprachaufnahmen vor laufender Kamera. > ei



Björn Eichelbaum, Ton-Ing. und BA Recording Arts, arbeitet seit 2002 freiberuflich in den Bereichen Sprache, Musik, Surround. Dazu testet er seit 2010 alle Workstations und viel weitere Hardware für die Digital Production.

#### Links

Mikrofonvergleich  
der besprochenen Geräte  
▷ [bit.ly/DP\\_audio\\_youtube](https://bit.ly/DP_audio_youtube)

Anzeige

# 34. Internationales KurzFilmFestival Hamburg

# 5. – 11. Juni 2018

[www.shortfilm.com](http://www.shortfilm.com)



KURZ FILM AGENTUR  
HAMBURG