# Fachbereich Populäre Musik Producing Composing Artist

Die moderne Vocal Produktion: der Mehrwert von Vocal Comping im Bezug auf die Entwicklung der Tontechnik.

Masterarbeit

eingereicht von:

Jonas Mengler

betreut duch:

Prof. Oliver Curdt, Dr. Heiko Wandler

# Inhaltsverzeichnis

	Inhaltsverzeichnis					
	Ehrer	Ehrenwörtliche Erklärung				
	Abkü	Abkürzungsverzeichnis				
	Vorw	ort	1			
1	Entw	icklung der Tontechnik	3			
	1.1	Die Anfänge der Tonaufnahmen	3			
	1.2	Der frühe Musikproduzent	7			
	1.3	Das elektrische Aufnahmeverfahren	11			
	1.4	Die Entwicklung von Tonstudios	16			
	1.5	Die Mehrspurtechnik	18			
	1.6	Die digitale Revolution	21			
2	Die n	noderne Vocal Produktion	26			
	2.1	Aufnahme	26			
	2.2	Vocal Comping	30			
		2.2.1 Prinzip	30			
		2.2.2 Gefühlsdefinition	32			
		2.2.3 Praktisches Beispiel	33			
		2.2.4 Grundvoraussetzungen	36			
		2.2.5 Prallel Comping vs. Post Comping	37			
		2.2.6 Schnitte	38			
		2.2.6.1 Das 4-Stufen-Modell	40			

		2.2.6.2 Pr	raktische Beispiele für fehlerhafte Schnitte	41
	2.2.7	Interview	mit Ken Lewis	45
2.3	Vocals im Mix			4
	2.3.1	1 Kompressor, Reverb & Co.		
	2.3.2	Dopplungen		
	2.3.3	Melodyne	e, VocALign & Autotune	52
2.4	Optimaler Workflow für die moderne Vocal Produktion			
	2.4.1	1 Logic Template		
	2.4.2	Shortcuts		58
	2.4.3	Equipmen	ıt	58
		2.4.3.1	Computer & Zubehör	58
		2.4.3.2	Arbeitsfläche	62
		2.4.3.3	Talkback-Pedal	63
		2.4.3.4	Kopfhörermixer	64
Fazit	und Au	ısblick in d	ie Zukunft	6
Bonu	ısmatieı	ral		V
Abbi	ldungsv	verzeichnis		V
Liter	aturverz	zeichnis		V

3

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass ich die Arbeit selbstständig und nur mit den angegebenen

Hilfsmitteln angefertigt habe und dass alle Stellen, die dem Wortlaut oder dem Sinne

nach anderen Werken entnommen sind, durch Angabe der Quellen als Entlehnungen

kenntlich gemacht worden sind.

\_\_\_\_\_

Ort, Datum

Unterschrift des Verfassers

III

# Abkürzungsverzeichnis

**Abb**. Abbildung

A/D Analog/Digital

**CC** Compact Cassette

**CD** Compact Disk

**cm** Zentimeter

**DAW** Digital Audio Workstation

**DAT** Digital Audio Tape

**dB** Dezibel

**Dr.** Doktor

**EQ** Equalizer

kHz Kilohertz

min Minuten

**mm** Millimeter

MP3 MPEG-1 Audio Layer III

ms Millisekunde(n)

**Prof.** Professor

**s** Sekunde(n)

#### Vorwort

Diese Arbeit beschäftigt sich mit der modernen Vocal Produktion. Hierbei geht es insbesondere um eine bestimmte Form der Nachbearbeitung von Gesangsaufnahmen, dem Vocal Comping. Dem resultierend wird diese Arbeit darlegen, welchen Mehrwert diese Technik in Bezug auf das Endergebnis der Produktion bietet. Arbeitsschritte wie die Vocal-Aufnahme und der Vocal-Mix werden dabei nicht ausgelassen, aber nur am Rande behandelt.

Ich beschäftige mich mit der Vocal Produktion seit meinem 12. Lebensjahr, zum heutigen Zeitpunkt sind das 15 Jahre. Nach einer klassischen Musikausbildung und dem Abitur am Landes-Musik-Gymnasium in Montabaur setzte ich meine Laufbahn an der Popakademie Baden-Württemberg fort, an der ich meinen Bachelor Of Arts absolvierte und mit dieser Abschluss-Arbeit nun den Master Of Arts anstrebe.

Auch wenn ich mit meiner Arbeit als Musikproduzent in verschiedenen Bereichen tätig bin (Beat Production, Score-Composing, Band Production, Mixing) hat sich in dieser Zeit für mich herauskristallisiert, dass die Arbeit mit Vocals im Studio zu den komplexesten und zugleich spannendsten Gebieten gehört.

Was mich motiviert meine Masterarbeit über dieses Thema zu schreiben ist die Wahrnehmung, dass die Kunst des Vocal Compings zwar Alltag in der heutigen Welt der Musikproduktion ist, wenngleich der wissenschaftliche Ansatz dieses Prinzips bislang nur rudimentär formuliert wurde. Zudem spielt für mich seit einigen Jahren die Optimierung des Workflows im Bereich der Vocal Produktion eine zentrale Rolle, weshalb ich beschlossen habe, mich auch diesem Thema in dieser Arbeit zu widmen.

Anmerkung: Wenn ich in dieser Arbeit von "dem Sänger" rede, dient dies lediglich der Vereinfachung. "Der Sänger" bezieht sich demzufolge gleichermaßen auf das weibliche Pendant "Die Sängerin".

Meine Danksagung geht in erster Linie an meine Eltern, die mich in all dem unterstützt haben, was mir dazu verholfen hat heute an diesem Punkt zu stehen. Des weiteren möchte ich der Popakademie Baden-Württemberg - stellvertretend dem künstlerischem Direktor

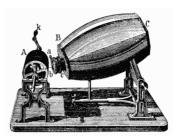
Udo Dahmen - für intensive, produktive und ereignisreiche sechs Jahre danken, die ich während meines Studiums an dieser Institution erfahren durfte. Nicht zuletzt richte ich meinen Dank an meine beiden Betreuer, Dr. Heiko Wandler & Prof. Oliver Curdt, sowie diejenigen, die mich während der Schreibphase dieser Arbeit moralisch unterstützt haben.

# 1. Entwicklung der Tontechnik

## 1.1 Die Anfänge der Tonaufnahmen

Wenn man sich die heutigen, für selbstverständlich gehaltenen Möglichkeiten der Aufnahme und Wiedergabe von Musik vor Augen hält, fällt es schwer sich eine Welt ohne Pro-Tools und MP3 vorzustellen. Um zu verstehen, mit welchen Möglichkeiten wir im Bereich der Tontechnik heute verwöhnt sind und um die Relevanz der modernen Vocal Produktion näher zu bringen, bedarf es zunächst einer Reise zurück in die Vergangenheit.

Den Grundstein für die allerersten Tonaufnahmen legte Édouard-Léon Scott de Martinville im Jahr 1857 mit der Erfindung des Phonautographen<sup>1</sup>. Der französische Tüftler, der sein Geld als Drucker und Korrekturleser wissenschaftlicher Werke verdiente, erfand und patentierte ein Gerät, mit dem es möglich war, Schall sichtbar zu machen. Dafür verwendete er einen beweglichen Trichter mit einer Membran am einen und einem Aufzeichnungsstift am anderen Ende. Dabei übertrugen



**Abb. 01** Édouard-Léon Scott de Martinville, Phonautograph,

sich die Schallschwingungen auf den Stift, der wiederum den Amplituden-Verlauf graphisch auf einen mit Ruß geschwärzten Papierstreifen zeichnete. Diese visualisierte Aufzeichnung des Schalls war allerdings nicht dafür gedacht, jemals wiedergegeben zu werden<sup>2</sup>. Die erste Tonaufnahme der Geschichte konnte im Nachhinein restauriert und auf den 09. April 1860 zurück datiert werden. Zu hören ist eine Sängerin mit einer 10-sekündigen Aufnahme des Folk-Songs "Au Claire De La Lune".<sup>3</sup>

<sup>&</sup>lt;sup>1</sup> Neologismus aus altgriechisch φωνή phoné "Τοn", αὐτός autós "selbst" und γράφειν gráphein "schreiben"; wörtlich also Tonselbstschreiber

<sup>&</sup>lt;sup>2</sup> erst 2008 konnte eine Tonaufzeichnung dieses Phonautographen rekonstruiert und hörbar gemacht werden

<sup>&</sup>lt;sup>3</sup> Vgl. Ronsen, 2008 - die restaurierte Aufnahme befindet sich auf dem beiliegenden Datenträger 1 unter *Entwicklung der Tontechnik* 

Als Weiterentwicklung des Phonautographen reichte der US-amerikanische Erfinder Thomas Alva Edison<sup>4</sup>, am 24. Dezember 1877 das Patent für seinen "Phonographen" ein<sup>5</sup>. Ähnlich wie bei dem Phonautographen wurde hier die Schwingung des Schalls durch einen Trichter eingefangen und über eine an die Membran eingespannte Walze in Zinnfolie eingegraben. Die Besonderheit bei diesem Gerät: Neben der Aufzeichnung von Schall war es erstmals auch möglich, das Aufgenommene über eine stumpfe Nadel in Verbindung mit einer Membran wiederzugeben. Der Schall wurde dabei über eine Schalldose verstärkt.

Edison erdachte eine Liste von möglichen Anwendung für seinen Phonographen:

- 1. Letter writing and all kinds of dictation without the aid of a stenographer.
- 2. Phonographic books, which will speak to blind people without effort on their part.
- 3. The teaching of elocution.
- 4. Reproduction of music.
- 5. The 'Family Record' a registry of sayings, reminiscences, etc., by members of the family in their own voices, and of the last words of dying persons.
- 6. Music-boxes and toys.
- 7. Clocks that should announce in articulate speech the time for going home, going to meals etc.
- 8. The preservation of languages by exact reproduction of the manner of pronouncing.
- 9. Educational purposes, such as preserving the explanations made by a teacher, so that the pupil can refer to them at any moment, and spelling or other lessons placed upon the phonograph for convenience in committing to memory.
- 10. Connection with the telephone, so as to make that instrument an auxiliary in the transmission of permanent and invaluable records, instead of being the recipient of momentary and fleeting communication.<sup>6</sup>

<sup>&</sup>lt;sup>4</sup> weitere Erfindungen: Glühbirne (1879), Elektrischer Stuhl (1888), Börsenkursanzeiger (1869) uvm.

<sup>&</sup>lt;sup>5</sup> Zeitgleich erfand Charles Cros ein Gerät mit derselben Funktion, ihm fehlten jedoch die finzanziellen Mittel, seine Erfindung, die er Paléophon nannte, zu patentieren und konstruieren zu lassen.

<sup>&</sup>lt;sup>6</sup> Vgl, Gelatt 1977, S.29



Abb. 02 Thomas Alva Edison mit seinem Phonographen

In dieser Liste möglicher Anwendungen finden sich unter den Punkten 4 und 6 zwar auch zwei Anwendungsmöglichkeiten, die die Speicherung und Wiedergabe von Musik betreffen, jedoch betrachtete Edison den Phonographen schwerpunktmäßig als ein Gerät zur Aufnahme, Wiedergabe und Übermittlung von Sprache. Im April 1878 gründete Edison zusammen mit fünf Eigentümern die Edison Speaking Phonograph Company, deren Aufgabe es war, sich um die Vermarktung des Geräts zu bemühen. Die Begeisterung für den Phonographen hielt sich allerdings in Grenzen, da der Nutzen dieses Geräts erheblich eingeschränkt war; die

Tonqualität war nicht ausreichend, so dass die Wiedergabe der Sprache teils unverständlich war. Das Gerät ließ sich zudem nur durch geschultes Personal bedienen, die Aufzeichnugsdauer war mit etwa einer Minute zu kurz, die sensible Zinnfolie eignete sich nur für wenige Wiedergaben und die Aufzeichnungen konnten aufgrund des Verschleißes weder dauerhaft aufbewahrt noch reproduziert werden.<sup>7</sup> Erst der spätere Einsatz von Wachswalzen anstelle der Zinnfolie ermöglichte schließlich eine wesentlich exaktere, differenziertere und längere Aufzeichnung. Zudem ließen sich die Wachswalzen erstmals durch ein Gussverfahren vervielfältigen. Allerdings stellte sich die Reproduzierbarkeit auch hier als durchaus problematisch heraus, wie Gelatt in folgendem Abschnitt beschreibt:

"The method of duplication involved "playing" a master cylinder on one mandrel and t ransferring the sound vibrations pantographically onto a wax blank revolving on an adjacent mandrel. Although it was slow and inadequate – not to be compared to the gramophone's stamping process – at least it allowed the mass production of records without imposing intolerable strain on the performers. Recording artists were obliged, nevertheless, to do many "repeat performances" before the recording horns, for one master cylinder would yield only a limited number of pantographed duplicates. Harry MacDonough, a veteran tenor of early cylinder days, recalled that Edison recording engineers in the late Nineties would make five master cylinders at each performance of a song and that each one of these masters would produce at least twenty-five duplicates before the original wax impressions wore out. A single "take" before five recording horns would thus yield 125 cylinders; but that would not begin to meet the demand for a fast-selling selection, and MacDonough remembered having recorded his popular songs again and again at the Edison studio in West Orange, New Jersey."8

<sup>&</sup>lt;sup>7</sup> Vgl. Gelatt 1977, S.30 f.

<sup>&</sup>lt;sup>8</sup> Vgl. Gelatt (1977), S. 81 f.

Dennoch brachte diese Weiterentwicklung, die Edison als den "Improved Phonograph" bezeichnete, schließlich den Startschuss für eine weltweit anlaufende Phonographenindustrie.

Anknüpfend an Edisons Erfindung meldete der deutschstämmige Amerikaner Emil Berliner im Jahr 1887 das "Grammophon" zum Patent an. Auch wenn es sich hierbei um ein reines Abspielgerät handelte, bot es zwei deutliche Vorteile gegenüber dem zu dieser Zeit konkurrierenden Phonographen: zum einen waren die neuen scheibenförmigen Tonträger, die Berliner "Schallplatten" nannte, platzsparender als die zuvor eingesetzten Walzen. Zum anderen konnten diese einfacher als Massenkopie und somit deutlich kostengünstiger produziert werden. Die Schallplatten hatten meist



Abb. 03 Das Grammphon

einen Durchmesser von 10 Zoll (25,4 Zentimeter) oder 12 Zoll (30,48 Zentimeter). 10-Zoll Platten boten etwas mehr als drei Minuten, 12 Zoll etwas mehr als vier Minuten Spielzeit pro Seite bei einer Abspielgeschwindigkeit von 70 Umdrehungen pro Minute. Die Kurbel des Grammphons wurde anfangs noch manuell per Hand betrieben, so konnte nur ein geübter Kurbler die Musik in der richtigen Tonhöhe hören. Den ausgelieferten Grammophonen war daher ein Warnzettel beigelegt: "Achtung! Mit 70 Umdrehungen kurbeln! Jede Umdrehung mehr oder weniger verändert die Tonart!"9. Die manuelle Inbetriebnahme der Kurbel des Grammophons wurde durch die Einführung des mechanischen Federwerksantrieb<sup>10</sup> im Jahr 1896 abgelöst.

Die ersten Erfolge der Schallplatte zeichneten sich schnell ab. Bis zum Jahr 1894 wurden etwa 1000 Grammophone und 25.000 dazugehörige Platten verkauft<sup>11</sup>. Im Oktober 1896 gab Berliner die Verwendung von Hartgummi als Plattenmatieral auf und ersetze die Substanz durch eine Pressmasse, die unter anderem aus dem Material Schellack<sup>12</sup> bestand. Neben einer größeren Strapazierfähigkeit der Neuentwicklung, die fortan den Namen "Schellackplatte" erhielt, bot diese dem Konsumenten zugleich eine enorme

<sup>&</sup>lt;sup>9</sup> Vgl. Haas, W. (1959), S. 35

<sup>&</sup>lt;sup>10</sup> Erfunden durch den Mechaniker Eldridge R. Johnson, 1867-1945

<sup>&</sup>lt;sup>11</sup> Tim Brooks, Richard Keith Spottswood: Blacks And The Birth of The Recording Industry 1890-1919. 2004, S. 35.

<sup>&</sup>lt;sup>12</sup> Schellack ist eine harzige Substanz, die aus Gummilack hergestellt wird.

Verbesserung der Klangqualität und Haltbarkeit. Das Grammophon legte schließlich den Grundstein für einen völlig neuen Wirtschaftszweig: die Tonträgerindustrie.

Am 6. November 1898 gründeten Emil und sein Bruder Joseph Berliner in Hannover die Deutsche Grammophon, eine Tochtergesellschaft des kurz zuvor in England gegründeten Mutterunternehmens Grammophone Company. Die Brüder setzen den Startschuss für die Massenproduktion der Schellackplatten. Lagen die Verkaufszahlen jener Platten im Jahr 1896 noch bei einer Stückzahl von 25.000, gingen bereits vier Jahre später insgesamt schon 700.000 Einheiten über den Ladentisch<sup>13</sup>. Zeitgleich wurden durch die am 19. Mai 1890 gegründete Fabrik Grammophon-Spielwaren-Fabrik Kämmer, Reinhardt & Co im thüringischen Waltershausen die ersten industriellen Grammophone hergestellt.



**Abb. 04** "His Masers Voice" Nipper

Die *Grammophone Company* wurde in "His Master's Voice" umbenannt. Der "Nipper", das Warenzeichen in Form eines Hundes, der in ein Grammophon schaut, wurde am 16. Juli 1900 registriert und fortan als Firmenlogo für die Schellackplatten verwendet. Im selben Jahr verlegte die *Deutsche Grammophon* ihren Verwaltungssitz nach Berlin und die Firma wurde in eine Aktiengesellschaft umgewandelt.

#### 1.2 Der frühe Musikproduzent

Berliners Geschäftspartner Fred Gaisberg reiste fortan mit einem mobilen Aufnahmegerät <sup>14</sup> rund um die Welt, um Künstler für die ersten kommerziellen Schallplatten-Aufnahmen zu

<sup>&</sup>lt;sup>13</sup> Tim Brooks, Richard Keith Spottswood: Blacks And The Birth of The Recording Industry 1890-1919. 2004, S. 40.

<sup>&</sup>lt;sup>14</sup> das Wort ,mobil' ist hier ein dehnbarer Begriff. Das Equipment für das Aufnahmegerät war in 6 großen Kisten verstaut, wie aus dem Artikel "The invention of the modern music star in a hotel bedroom in Milan hervorgeht". (http://soundofthehound.com/2011/01/16/the-invention-of-the-modern-music-star-in-a-hotel-bedroom-in-milan/)

gewinnen. Seine Aufnahmen fanden meist in angemieteten Hotelräumen statt. Folglich war Gaisberg sowohl erster Schallplattenproduzenten als auch erster A&R<sup>15</sup>.

"Gaisberg's attitude to recording was to produce in the studio some kind of snapshot of the kind of Performance each artist would normally give in public venues."<sup>16</sup>

Am 11. April 1902 stieß in Mailand auf den Opernsänger Enrico Caruso. In einem Hotelzimmer richtete Gaisberg sein mobiles Aufnahmestudio ein. Die Aufnahme erfolgte über einen Trichter, der die Schwingungen per Nadel direkt auf die Wachswalzen übertrug. Caruso zeichnete eine Selbst-Karrikatur dieser Aufnahmesituation (siehe Abb. 05).

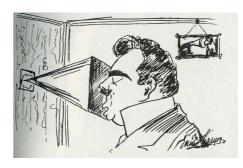


Abb. 05 Enrico Caruso während einer Aufnahme, 1902, Selbst-Karrikatur

Gaisberg bot Caruso 100 Pfund für die Aufnahme von

zehn Arien, dies stieß im Londoner Büro mit den Worten "FEE EXORBITANT FORBID YOU TO RECORD" <sup>17</sup> allerdings auf keine positive Resonanz. Gaisberg hingegen ignorierte die Anweisung. Anscheinend schien er zu ahnen, welches Erfolgskonzept sich hinter diesen Aufnahmen verbergen wird. Die *Grammophone Company* hätte 2000 Einheiten des Tonträgers verkaufen müssen, um die investierten 100 Pfund einzuspielen. Da zu der Zeit nur wenige Menschen über ein eigenes Grammophon verfügten und der Absatzmarkt der Schellackplatten demzufolge sehr gering war, erschienen 2000 Verkäufe tatsächlich eher unwahrscheinlich. Allerdings: Die 10 Arien des Enrico Caruso erreichten eine Verkaufszahl von 300.000 Einheiten<sup>18</sup>. So erlangten Caruso und die Grammophon nicht nur einen exorbitanten finanziellen Reichtum, sondern stießen zugleich auf weltweiten Erfolg; mit dem Opernsänger Enrico Caruso wurde der erste Welt-Star geboren und Gaisberg genoss den Ruhm, jenen Star entdeckt zu haben. Fred Gaisberg setze seine Reise und damit seinen Siegeszug mit Aufnahmen von Sängern aus der Opernszene rund um den Globus fort.

<sup>&</sup>lt;sup>15</sup> A&R steht für Artists and Repertoire und ist ein Begriff aus der Musikwirtschaft. Ein A&R-Manager sucht im Auftrag einer Plattenfirma nach neuen Talenten.

<sup>&</sup>lt;sup>16</sup> Vgl. Beadle (1993), S. 27

<sup>&</sup>lt;sup>17</sup> Vgl. Artikel "The invention of the modern music star in a hotel bedroom in Milan"

<sup>&</sup>lt;sup>18</sup> Vgl. Artikel "The invention of the modern music star in a hotel bedroom in Milan"

Mit steigenden Bedürfnissen nach professionellen Aufnahmen stieg die Nachfrage nach Personen, die in der Lage waren, diese Aufnahmen optimal zu gestalten. Mit dem neuerschaffenen Arbeitsfeld des Tontechnikers ging ein spezifisches Anforderungsprofil einher. Vom Tontechniker wurde die Ausnutzung und die Beherrschung aller Aspekte verlangt, die den Klang der Aufnahme bestimmten, um die Tonquelle an die Bedingungen des Aufnahmegeräts anzupassen. Dazu zählten vor allem

- die akustischen Eigenschaften des Aufnahmeraums,
- die Anzahl und Positionierung der beteiligen Musiker vor dem Trichter,
- das eventuelle Umarbeiten des Notenmaterials zur Optimierung der Klangeigenschaft, sowie
- das Beherrschen der gesamten technischen Apparatur<sup>19</sup>.

All diese Anforderungen basierten in der Regel auf eigenen Experimenten. Die gesammelten Erfahrungen stellten für die Tonträgerfirma ein entscheidendes Kapital dar. So erklärt sich, dass Aufnahmetechniker zu jener Zeit durch eine besondere Vertragsklausel zur "strengen Geheimhaltung der Aufnahmepraktiken"<sup>20</sup> verpflichtet wurden unter Androhung einer hohen Konventionalstrafe bei Zuwiderhandlung.

Von einem Tontechniker verlangte man jedoch nicht nur die Beherrschung der Technik und der Akustik, sondern auch eine musikalische Begabung, wie aus einem Anforderungsprofil für einen Tontechniker - verfasst vom deutschen Komponist und Wissenschaftler Max Chop im Jahr 1911 - hervorgeht:

"Für die grammophone Kunst ist es von fundamentaler Bedeutung, dass der Techniker auch ein Musiker sei (was durchaus nicht immer zutraf), d.h. nicht bloss ein Mann, der korrekte Töne von falschen zu unterscheiden vermag, vielmehr auch über die inneren Qualitäten dieser Töne sich durchaus klar bleibt, er die Geheimnisse des Zusammenklangs mit all den vielfältigen Mischungen durchschaut und dort, wo Differenzen kleiner Art zwischen Darbietung und Reproduktion bestehen, auf deren Ausmerzung hinarbeitet, alle nicht hingehörigen Beimischungen

-

<sup>&</sup>lt;sup>19</sup> Vgl. Gauß (2009), S. 176

<sup>&</sup>lt;sup>20</sup> Zitat O. Multhaupt: Die praktische Aufnahmetechnik, PZ (Phonographische Zeitung) 10.1909, Nr. 28 S. 666 f.

beseitigt, dabei von absolut musikalischen Voraussetzungen ausgeht und mit seiner Ästhetik vollkommen ,up to date' ist."<sup>21</sup>

Eine beispielhafte Situation für die Arbeit eines Tontechnikers schildert die Sängerin Frieda Hempel, von 1907 bis 1912 an der Berliner Hofoper engagiert, in einer Verfahrensweise zum Ausgleich der Dynamik während der Aufnahmeprozedur:

"Ich wurde in ein verhältnismäßig kleines Zimmer geführt. Vor mir stand ein trichterförmiges Horn, in das ich singen sollte. Neben ihm waren Apparaturen, an denen mehrere Männer aufgeregt hantierten. Irgendeiner der Männer knuffte mich in den Rücken, es war das Zeichen, dass ich singen sollte, und ich sang. Wenn meine Partitur kräftigere Töne verlangte, zog mich einer der Techniker am Rock, damit ich zurücktrat. Wenn ich piano zu singen hatte, schob er mich mit entsprechender Energie an den Trichter heran. Bei diesem handgreiflichen Verfahren die musikalische Kontinuität zu wahren, fiel nicht leicht. Die Aufnahme mußten sehr oft wiederholt werden, weil die Stimme nicht gleichmäßig war. Und wenn eine Aufnahme wirklich gelückte, dann zerbrach wohlmöglich die Platte"<sup>22</sup>

Daraus wird deutlich, welche Experimentierfreudigkeit dem Tontechniker zu jener Zeit abverlangt wurde. Die Aufnahmetechnik um 1908 arbeitete "sichtlich auf die Verfeinerung der Musik" hin.<sup>23</sup>



**Abb. 06** Aufnahme der Berliner Philharmoniker unter Arthut Nikisch, 1913

Die erste Aufnahme eines vollbesetzten Orchesters, den Berliner Philharmoniker unter Arthur Nikisch wurde für die *Deutsche Grammophon Gesellschaft* getätigt und geht auf das Jahr 1913 zurück. Wie auf der Abb. 06 zu erkennen ist wird auch hier auf die selbe Aufnahmemethode zurückgegriffen, die einst Fred Gaisberg in seinen Hotelzimmern angewendet hat; ein einziger Trichter bespielt eine Wachswalze oder eine

<sup>&</sup>lt;sup>21</sup> Zitat Max Chop: Die künstlerische Jahresbilanz der sprechmaschinellen Produktion, PZ 12.1911, Nr. 1, S. 5

<sup>&</sup>lt;sup>22</sup> Zitat nach Dieter Glatzer u. Ruth Glatzer: Berliner Leben 1900-1914: Eine historische Reportage aus Erinnerungen und Berichten (DEB) Berlin, 1986, S. 592

<sup>&</sup>lt;sup>23</sup> Vgl. Noricus: Neue Erfahrungen in der Aufnahmetechnik, PZ 10.1909, Nr. 16, S. 428

Schellackplatte<sup>24</sup>. Soll ein Instrument beziehungsweise eine Instrumentengruppe auf der Aufnahme lauter erscheinen, so müssen diese sich näher an dem Trichter positionieren. Im Fall von Caruso war seine Stimme näher am Trichter als das im Hintergrund spielende Klavier, im Falle einer Orchesteraufnahme stehen die von Natur aus lauteren Bläsergruppen weiter vom Trichter entfernt als die Streichergruppen.

Die Möglichkeit, verschiedene Instrumentengruppen über verschiedene Trichter aufzunehmen erläuterten erstmals die beiden Autoren Roger Beardsley und Daniel Leech-Wilkinson in ihrem Artikel "A Brief History of Recording to ca. 1950" über eine typische Aufnahme Session im Jahr 1920:

"The recording medium was a form of hard wax, prepared at the factory in vats, filtered to keep the mixture as smooth as possible and, after cooling and hardening in circular moulds, turned on a lathe in order to produce a smooth surface. The resulting wax blanks were packed in cases and sent out to the studios. For recording, the wax was placed on the turntable of the recording machine, where it rotated in theory (though by no means always in practice) at 78 revolutions per minute (rpm). The whole turntable assembly moved sideways beneath the cutting head so that a groove was cut in the wax from the outer edge of the wax disc towards the centre. Because it was attached directly to the horn, the cutting assembly was fixed. The stylus was moved by a (usually glass) diaphragm, and, depending on how efficient the connections were, pressure was transmitted more or less accurately. Sound was transmitted to the other side of the diaphragm through tubular connections leading from the narrow end of the recording horn, while the horn itself captured and focussed as much of the performers' sound output as possible. Horns tended to have resonances of their own, which were damped as far as possible by wrapping tape around them. Multiple horns could be used to capture sound from a larger group of performers or from different parts of a piano, for example, and these were connected up via Y-shaped metal connectors joined to the horn with rubber tubing. The horns were suspended from (or occasionally supported on) stands to minimize strain on the cutting mechanism."25

#### 1.3 Das elektrische Aufnahmeverfahren

Mit zunehmendem Perfektionismus innerhalb der Aufnahmeprozesse wurde der Wunsch stärker, auch Tonaufnahmen in besserer Qualität zu produzieren. Als im 1924 das

<sup>&</sup>lt;sup>24</sup> Dies lässt sich nicht genau sagen, da zu dieser Zeit noch beide Systeme existierten. Erst im Sommer 1929 wurden die Wachswalzen als Aufnahmemedium endgültig eingestellt.

<sup>&</sup>lt;sup>25</sup> Auszug auf dem Artikel "A Brief History of Recording to ca. 1950", erschienen auf der Webseite http://www.charm.rhul.ac.uk/history/p20\_4\_1.html (Stand: 02. Juli 2014), Erscheinungsjahr unbekannt

elektrische Aufnahmeverfahren und wenig später auch das elektrische Wiedergabeverfahren ermöglicht wurden, verbesserte sich die Tonqualität der Schallplatten gewaltig.

Ein fundamentaler Baustein für das elektrische Aufnahmeverfahren war die Verwendung eines Mikrofons. Bereits 1861 führte der Erfinder Philip Reis der Physikalischen Gesellschaft zu Frankfurt am Main eine Apparatur vor, die mit Hilfe eines Schweinedarms Schall in elektrische Impulse umwandelte - die Basis für das Mikrofon. Georg Neumann entwickelte im Jahr 1923 das Mikrofon anhand der Erkenntnisse von Reis weiter, wodurch die Klangqualität besonders bei tiefen Frequenzen stark verbessert wurde. Der Durchbruch gelang ihm jedoch mit der Entwicklung



**Abb. 07** 1928 Neumann CMV3 Kondensatormikrofon

des Niederfrequenz (NF)-Kondensatormikrofons. Membran und Gegenelektrode bildeten hier einen Kondensator, der auf eine Gleichspannung aufgeladen wird; durch die Membranbewegung änderte sich die Kondensatorkapazität. Aus dieser Wurde das Signal gewonnen. Dieses Wandlerprinzip war der Schallaufzeichnungstechnik seiner Zeit qualitativ weit voraus und ist bei Mikrofonen höchster Qualität noch heute Standard.

Anstelle des Schalltrichters verwendete man bei der Tonaufnahme nun das Mikrofon, das die Schallschwingungen in elektrischen Strom umwandelte. Die Führung des Schneidestichels durch die Membran wurde dabei über Mikrofon und Verstärker von einem Elektromagneten übernommen, zwischen dessen Polen sich der Schneidestichel im Rhythmus der Musik oder Sprache hin und her bewegte. Der elektrische Röhrenverstärker war das entscheidende Bindeglied in der Aufnahmekette, denn er musste die schwachen Signale des Mikrofons so verstärken, dass damit der elektromagnetische Schneidestichel bewegt werden konnte.

Die Vorteile der elektrischen Aufnahmen waren

- eine höhere Lautstärke,
- geringeres Rauschen, sowie
- erweiterter und natürlicherer Frequenzgang.

Im August 1925 kam die amerikanische Columbia mit der ersten elektrisch aufgenommenen Schallplatte auf den Markt. Am 01. März 1926 veröffentlichte die Homophone Company

auf ihrer Marke *Homocord* die erste in Deutschland elektrisch aufgenommene Schallplatte.<sup>26</sup>



Abb. 08 Victor Orchestra bei einer Aufnahme mit einem Tricher. 1925



Abb. 09 Victor Orchestra bei einer Aufnahme mit einem Mikrofon, 1925

Allerdings verloren mit dem elektrischen Aufnahmeverfahren die bis dahin im Experiment gewonnenen Erfahrungen ihre Gültigkeit zugunsten mathematisch berechenbarer Elektrophysik.<sup>27</sup> Diese Tatsache musste auf die Tontechniker, die sich die vorherigen Jahre zu Experten im Bereich der Aufnahme entwickelt hatten, wie ein Schock gewirkt haben. So schrieb der Tontechnik-Pionier Fred Gaisberg, nachdem er im Herbst 1924 eine Aufnahme des Elektrotechnikunternehmens *Western Electric* gehört hatte:

"Was die Western Electric sozusagen als Nebenprodukt ihrer Forschung auf dem Telefon-Nachrichtensektor erreicht hatte, ließ in meiner Welt eine Mine hochgehen. Meine Kollegen, allein versiert in der einfachen akustik-mechanischen Aufnahmemethode, mußten neu anfangen und dazu Elektrotechnik studieren. Mit Bestürzung mußten sie sehen, wie junge Elektrotechniker ihnen ihre Arbeit abnahmen, auf die sie sich in langer Lehrzeit eingearbeitet hatten. Nur einige meiner alten Kollegen meisterten den Übergang."<sup>28</sup>

<sup>&</sup>lt;sup>26</sup> Vgl. Online-Artikel "DIE GESCHICHTE DER TONAUFZEICHNUNG" http://rdtev.de/index.php? option=com\_content&view=article&id=18:interessantes-zur-geschichte-dertonaufzeichnung&catid=11&Itemid=112 (Stand: 07.07.2014)

<sup>&</sup>lt;sup>27</sup> Vgl. Gauß (2009)

<sup>&</sup>lt;sup>28</sup> Fred Gaisberg, zit. nach Bruch: Tonwalze (1983), S. 66 f.

Offensichtlich klar, dass keine Schallplattenfirma ohne das elektrische Aufnahmeverfahren in Zukunft noch konkurrenzfähig sein würde. So entstand von nun an der Bedarf an Aufnahmetechnikern mit neuen Kompetenzen. Ingenieure, die Erfahrung bei der Entwicklung der elektronischen Tonverstärker und Lautsprecher gesammelt hatten, übernahmen die Position früherer Aufnahmeleiter. Die Entwicklung der Aufnahmeverfahren verlagerte sich in die Laboratorien der Elektroindustrie, die zu einem bedeutenden Teil der Schallplatteninsdustrie wurde.<sup>29</sup> Siegfried Goslich, Leiter der Arbeitsgemeinschaft "Rundfunk und Tonfilm" beim Musikwissenschaftlichen Institut der Berliner Universität prägte im Jahr 1933 den Begriff des "Musikingenierus". Laut seiner Argumentation hätten Technik und Musikwissenschaft zusammengefunden und es bestehe ihre völlige Verschmelzung in dem Grenzgebiet einer "technischen Musikwissenschaft bevor". Er erläuterte weiter: "Als wirklicher Fachmann wird in Zukunft derjenige zu gelten haben, der eine wissenschaftliche Vorbildung eben auf dem Grenzgebiet der ,technischen Musikwissenschaft' mit praktischer Erfahrung vereint."30.

Die elektrische Aufnahme der Schallplatten war fortan allgemeiner Stand der Technik, wenn gleich die Wiedergabe der Platten immer noch über Grammophone erfolgten.



Abb. 10 Magnetophon K1

Mit der Präsentation des "Magnetophon K1" legte die Firma TELEFUNKEN auf der Funkausstellung 1935 in Berlin, den Grundstein der modernen Tonaufzeichnung. Ein Speichermedium war geschaffen worden, das nach dem Prinzip der 'Merkfähigkeit' magnetisch geladener Teilchen funktionierte. Als Trägerbänder fungierten anfangs eisenoxydbeschichtete Papierstreifen, die der Signalstärke entsprechend durch Induktionsströme magnetisiert werden

konnten. <sup>31</sup> Das Gerät besaß den neuen Dreimotorenantrieb<sup>32</sup>, einen schnellen Rücklauf und

<sup>&</sup>lt;sup>29</sup> Vgl. Gauß (2009), S. 178

<sup>&</sup>lt;sup>30</sup> Edb., Zit. S 684, zitiert nach Gauß (2009) S. 179

<sup>31</sup> Vgl. Wernicke (1998)

<sup>&</sup>lt;sup>32</sup> Der Dreimotorenantrieb wurde 1935 von der BASF entwickelt. Er erfügte über einen Antriebsmotor für die Wiedergabe und je einen für das Vorspulen und für das Zurückspulen.

Drucktastensteuerung. Das K1 bestand aus drei schweren Koffern, dem Laufwerk, dem Verstärker und dem Lautsprecher. Die technischen Daten von K1 waren:

- Bandgeschwindigkeit 100 cm/sek.
- 25 cm Spulendurchmesser,
- Bandwickel 1000 m,
- Spieldauer 20 min.,
- Frequenzgang 50 Hz bis 5 kHz,
- Geräusch-Spannungs-Abstand ca. 35 dB und Bandbreite 6,5 mm (heute 6,3). 33

Die Arbeit mit den magnetischen Tonbändern brachte zudem eine revolutionäre Eigenschaft mit sich: Erstmals in der Geschichte der Tonaufzeichnung war es möglich, bestimmte Teile einer Aufnahme auszuschneiden, einzufügen, zu verschieben oder sogar verschiedene Aufnahmen zu kombinieren.

Bei dem Magnetbandschnitt ging man folgendermaßen vor:

- 1. Schnittstellen wurden durch Spulen von Hand gesucht und mit Filzstift markiert. (bei beiden Bändern, die zusammengefügt werden sollen)
- 2. Das Band wurde mit eingebauter Schneidevorrichtung oder Bandschere (unmagnetisch) geschnitten.
- 3. Die Bänder wurden mit einem ca. 4 cm langen Klebestreifen verbunden.
- 4. Überblendungen konnten durch diagonalen Schnitt des Bandes realisiert werden.
- 5. Nur durch handwerkliches Geschick ließen sich längere Blenden erstellen. Maximale Länge ca. 100ms.<sup>34</sup>

<sup>&</sup>lt;sup>33</sup> Vgl. Online-Artikel: <a href="http://www.radiomuseum.org/r/aeg\_magnetophon.html">http://www.radiomuseum.org/r/aeg\_magnetophon.html</a> (Stand 07.07.2014)

<sup>&</sup>lt;sup>34</sup> Vgl. Hartmann (2009 oder 2010)

1.4 Entwicklung von Tonstudios

Mit stetiger Weiterentwicklung der Aufnahmetechniken und der Abspielgeräte schossen

Plattenfirmen und damit zusammenhängend Musikstudios aus dem Boden, etwa am 12.

November 1931 die Abbey Road Studios in London. Unter der Adresse 3 Abbey Road, St.

John's Wood (City of Westminster), London NW8 9AY, verbarg sich zunächst ein 1830

errichtetes Wohngebäude im georgianischen Baustil, das am 3. Dezember 1929 von The

Gramophone Company Ltd. für 16.500 Pfund erworben und in ein Tonstudio umgebaut

wurde. Die Umbaukosten betrugen mit 100.000 Pfund ein Vielfaches des Kaufpreises. 35

Alle klassischen symphonischen Werke der EMI entstanden fortan hier in den Studios.<sup>36</sup>

1929 spricht die BBC in ihrem Hand Book erstmals von "Mixing Studios"<sup>37</sup> und erklärt

den noch in Anführungszeichen gesetzten Begriff so:

In längeren Rundfunkproduktionen wie zum Beispiel Hörspielen, die damals live

aufgeführt wurden, gab es zwei Typen von Klangquellen – die Sprechstimmen und die

Geräusche. Ursprünglich waren beide in einem Raum untergebracht, aber die Hörer

beschwerten sich, bei lauten Effektgeräuschen der Erzählung nicht mehr folgen zu können.

Als Konsequenz lagerte der Londoner Sender die "Noise Effects" (Gewitter durch große

Metallfolien, Pferdegalopp durch Stein auf Stein und so weiter) in einen gesonderten Raum

aus; die Effektemacher hörten über Kopfhörer mit, was im Sprecherraum geschah.38 So

schrieb Anon in seinem Bericht für das BBC Handbook:

"The sounds from both studios were transmuted by lines to a central switchboard

under the control of the senior producer and he was thus enabled to "mix" them in

the exact quantities required."39

Anon spricht dabei auch erstmals von dem "Switchboard" (Schaltpult), welches den

Vorläufer des Mischpults darstellt. Das Konzept des Aufnahmeprinzips war so erfolgreich,

 $^{35}$  Vgl. Martland (1997), S. 248 f.

 $^{36}$  Vgl. Martland (1997), S 81ff.

<sup>37</sup> Zit. Anon: BBC Handbook (1929) S. 180

<sup>38</sup> Vgl. Anon: BBC Handbook (1929) S. 180

<sup>39</sup> Zit. Anon: BBC Handbook (1929) S. 180

16

dass der Sender große Produktionen Ende der 1920er Jahre mit mehr als drei Studios durchführte. In einem saß ein Orchester, in einem anderen eine Band; auch die Schauspieler wurden in Gruppen getrennt, um verschiedene Akustiken herzustellen.

Der Aufbau der Tonstudios erfolgte zu dieser Zeit nach einem nach einem ganz bestimmten und im Grunde einheitlichen Konzept.

#### Regieraum

Dieser Raum fungiert als eine Art Zentrale für das Tonstudio. Hier befinden sich in der Regel Tontechniker, Tonmeister oder Aufnahmeleiter. Herzstück dieses Raums ist das Mischpult, in dem alle Mikrofonsignale beziehungsweise Bandmaschinensignale, deren Wiedergabe in laufende Aufnahmen eingeschleust wurden, zusammenlaufen.



Abb. 11 Abbey Road Regieraum, 1931

#### Aufnahmeraum



**Abb. 12** Abbey Road Studios Aufnahmeraum, 1931

Im Aufnahmeraum befinden sich ein oder mehrere Musiker, die meist durch eine Glasscheibe von dem Regieraum getrennt sind, so dass ein gegenseitiger Blickkontakt besteht. Im Aufnahmeraum sind ein oder mehrere Mikrofone positioniert, deren Signale in das Mischpult des Regieraums verlaufen.

#### Maschinenraum

Der Maschinenraum, der in der Regel angegliedert an den Regieraum seinen Platz findet, beherbergt technische Geräte wie die Bandmaschine oder Endstufen der



Abb. 13 Tonbandgeräte im Maschinenraum

Lautsprecher-Verstärker. Der Vorteil der Stationierung der Geräte in diesem Raum ist, dass ein störender Geräuschpegel im Regieraum vermieden wird. Von entscheidender Relevanz ist dieser Raum allerdings nicht, da die Geräte auch direkt im Regieraum platziert werden können.

### 1.5 Die Mehrspurtechnik

Obwohl die magnetische Schallaufzeichnung viele Vorteile mit sich brachte, konnte sie in puncto Sound-Qualität mit der damaligen Schallplatte nicht konkurrieren. Erst durch das Prinzip der Hochfrequenz-Vormagnetisierung, auf das der Einfinder Dr. Walter Weber im Jahr 1940 rein zufällig gestoßen war, wurde der Rauschpegel des Magnettonbands erheblich reduziert. Dieser Vorstoß katapultierte die magnetische Schallaufzeichnung an die Spitze der damaligen Aufzeichnugsmedien.<sup>40</sup>

Seitdem wurde auch das Magnetophon stetig weiterentwickelt. War in der Anfangszeit dieser Entwicklung das gesamte Magnetband nur einseitig bespielbar (Vollspurtechnik oder Mono-Technik), erreichten im Jahr 1943 die die ersten sogenannten "Halbspurgeräte" ihre Marktreife. Der Unterschied zu den bisherigen Geräten war, dass hier zunächst die Hälfte der Bandbreite bespielt, dann die Spulen am Bandende umgelegt und schliesslich die andere Hälfte bespielt wurde. So erhielt man die doppelte Spielzeit bei gleicher Bandlänge. Zudem legte die Halbspurtechnik den Grundstein für ein neues, revolutionäres Produktions-Werkzeug: die Stereophonie.

Als Stereophonie wird die Technik bezeichnet, die mit Hilfe von zwei Schallquellen einen räumlichen Schalleindruck beim natürlichen Hören erzeugen.<sup>41</sup> Um diesen Höreindruck zu erreichen wurden für die Aufnahmen zwei Mikrofone in einer bestimmten Anordnung verwendet, die jeweils eine der beiden

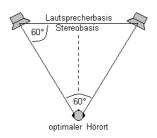


Abb. 14 Schaubild Stereophonie

<sup>&</sup>lt;sup>40</sup> Friesecke (2007), S. 596

<sup>&</sup>lt;sup>41</sup> Vgl. http://fakten-uber.de/stereofonie (Stand: 10.07.2014)

Halbspuren gleichzeitig bespielten. Die Wiedergabe erfolgte über zwei Lautsprecher, die wiederum jeweils einer der Halbspuren zugewiesen waren. Spricht man seither also von einem Stereoginal, so beinhaltet dieses zwei Mono-Signale beziehungsweise Spuren.

Walt Disney brachte seinen Zeichentrickfilm Fantasia (1940) in "Fantasound", einem frühen stereofonischen Tonverfahren heraus, das zur damaligen Zeit aber nur die wenigsten Kinos wiedergeben konnten.<sup>42</sup> Sidney Frey, Chef des Plattenlabels Audio Fidelity Records, brachte 1957 die erste Stereo-Schallplatte heraus: Auf der linken Hörseite waren Eisenbahngeräusche zu hören, auf der rechten Dixieland Jazz mit den Dukes of Dixieland<sup>43</sup>.

Die Weiterentwicklung zur Vierspur-Technik brachte darauf hin die Möglichkeit mit sich, zwei Stereo Quellen beziehungsweise vier Mono Quellen gleichzeitig aufzunehmen. Die Möglichkeit der 4-spurigen Bandaufnahme ließ die damaligen Produzenten zu neuer Experimentierfreudigkeit verleiten, wie Kevauver in seinem "Audio Recording Handbook" beschreibt:

"Once magnetic tape became the standard studio recording medium, it was only a matter of time before musicians began adding accompaniments to their recordings by playing along with a previously recorded tape. Both the new and the previously taped performance would be mixed together and recorded onto a second tape recorder. This technique became known as overdubbing."<sup>44</sup>

Seitdem sich der Einsatz von Bandmaschinen im Tonstudio zum Standard etablierte, eröffnete sich also die Möglichkeit des Overdubbings. Diese Technik lässt sich folgendermaßen beschreiben: Ein Orchester soll zusammen mit Gesang für eine Stereo-Produktion aufgenommen werden. Im ersten Schritt wird das gesamte Orchester auf diese Stereo-Spur (beziehungsweise zwei Mono-Spuren) auf die "Bandmaschine 1" aufgenommen. Im zweiten Schritt nimmt der Sänger auf, er hört dabei das bereits aufgenommene Orchester, das von der "Bandmaschine 1" abgespielt wird. Über das

<sup>&</sup>lt;sup>42</sup> Vgl. Kerins (2010), S. 23

<sup>&</sup>lt;sup>43</sup> Vgl. Russell Sanjek American Popular Music and Its Business: From 1900 to 1984 1988, S. 360

<sup>&</sup>lt;sup>44</sup> Zit. Kefauver (2001), S. 503 f.

Mischpult balanciert der Toningenieur das Verhältnis der Stimme zum Orchester. Dieser Signalmix wird schliesslich auf eine zweite zweispurige Bandmaschine aufgenommen.

Der Vorteil, den dieses System bietet, liegt auf der Hand: Soll eine Stelle im Gesang wiederholt werden, muss nur der Sänger und nicht zusätzlich das gesamte Orchester diese Stelle wiederholen. Der Gesang ist also unabhängig vom begleitenden "Playback". <u>Diese Technik bietet also erstmal eine Flexibilität im Bezug auf die Vocal-Aufnahme, die zwar unter Umständen zu mehr Arbeit, dafür aber gleichzeitig zu einer wachsenden Perfektion führen kann.</u>

Die Weiterentwicklung der Bandmaschine bis hin zu 24-Spur-Geräten ermöglichten schliesslich nicht nur die gleichzeitige Aufnahme 24 einzelner Tonquellen, sondern auch das nachträgliche Eingreifen in jede Spur (Stichwort: Bandmaschinen-Schnitt) sowie das nachträgliche Abmischen dieser Spuren. Mehrspurige Bandmaschine etablierten sich zum Standard im Tonstudio-Bereich. Erwähnenswert ist hier beispielsweise das Modell *Studer A800* (siehe Abbildung 15).



**Abb. 15** Bandmaschine Studer A800

Parallel zur kontinuierlichen Verbesserung der Aufnahmetechnik arbeiteten Wissenschaftler stetig an der Weiterentwicklung der Schallplatte, die zu diesem Zeitpunkt immer noch als Wiedergabemedium für den Massenmarkt diente. Am 21. Juni 1948 gelang mit der Erfindung der Vinyl-Schallplatte durch den ungarisch-amerikanischen Physiker Dr. Peter Carl Goldmark (1906-1977) schließlich der Durchbruch auf dem Tonträgermarkt. Im gleichen Jahr stellte die amerikanische Plattenfirma Columbia die erste Kunststoffplatte mit Mikrorille und 33 1/3 Umdrehungen pro Minute (UpM) vor.

Die ersten Vinyls hatten Durchmesser von 30 cm und wurden als "long playing record" oder in der deutschen Übersetzung als Langspielplatte (LP) bezeichnet. Die Firma RCA Viktor führt 1949 die Single-Platte ein, die einen Durchmesser von 17,5 cm bei 45 UpM aufwies. Mit der Erfindung der Vinyl-Schallplatte wurde auch das Ende der Schellackplatte eingeleitet, die weitaus empfindlicher und teurer waren. Als letztes stellte EMI die Produktion im Jahr 1958 ein. Des weiteren bot die Vinyl-Schallplatte erstmal die Möglichkeit der Wiedergabe im Stereo-Format.

Nur wenige Jahre später leitet die auf einem Magnetband basierende Compact-Cassette ein neues Zeitalter der Tontechnik ein. Die Compact-Cassette ist ein Tonband, das zur einfacheren Handhabung von einem Kunststoffgehäuse

umgeben ist. Sie wurde 1963 von dem holländischen Konzern *Philips* patentiert und weltweit auf die auch noch heutzutage gebräuchliche Form genormt. Die Magnetbänder können in Mono oder Stereo von beiden Seiten bespielt werden. Die Kassette ist 10 cm lang, 6,8 cm breit und 0,8 cm dick. Im selben Jahr brachte *Philips* auch den ersten Kassetten Recorder auf den Markt. Mit diesem Gerät war es erstmals möglich, eigene Aufnahmen



Abb. 16 Kassetten Recorder der Firma Philips Typ EL 3302

für den Heimgebrauch zu tätigen. Insbesondere Jugendliche waren von der Möglichkeit begeistert, auf billige und einfache Art ihre Lieblingshits aus dem Radio aufnehmen zu können, woraufhin die Musikindustrie existenzbedrohende Umsatzrückgänge befürchtete und mit Kampagnen wie "Home Taping Is Killing Music" reagierte.

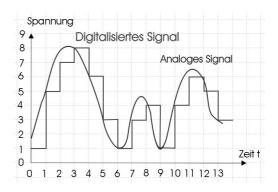
Mit dem technischen Fortschritt der Magnetbänder ging ein entsprechender Fortschritt in der Klangqualität der Compact-Cassette einher. Neben der Weiterentwicklung von Eisenoxidbändern sorgte auch die Einführung von Chromdioxid- und später Reineisen-Beschichtungen für einen wahrnehmbaren Qualitätssprung. Mit der Erfindung des "Walkman" im Jahr 1979 gelang es der Firma Sony erstmals ein Gerät zu schaffen, das den mobilen Konsum von Musik ermöglichte. Auch wenn die Compact-Cassette einen gewaltigen Fortschritt im Bereich der Audio-Aufnahme und Wiedergabe darstellte, sollte diese Technik schon bald von der digitalen Revolution eingeholt werden.

#### 1.6 Die digitale Revolution

 $<sup>^{45}</sup>$  Home Taping Is Killing Music war eine Kampagne der British Phonographic Industry. Sie wurde 1980 ins Leben gerufen.

Vgl. http://www.musikexpress.de/das-archiv/article156696/home-taping-is-killing-music-and-its-illegal.html (Stand: 10.07.2014)

Mit digitalen Aufzeichnungsmöglichkeiten auf Magnettonbändern wurde schon in den 70ern experimentiert. Der Analog-Digital Wandler<sup>46</sup> (auch A/D-Wandler oder ADC für *Analog-Digital-Converter*) ist hierbei das Schlüssel-Werkzeug für eine digitale Aufnahme. Analoge Signale müssen hierzu erst in digitale und diese anschließend wieder in analoge Signale umgewandelt werden.



**Abb. 17** Umwandlung eines analogen Signals in ein digitales Signal

Bei der Digitalisierung wird die Schallwelle in

einzelne Stufen (sogenannte Bits) zerlegt und deren Werte im Binärsystem<sup>47</sup> gespeichert. Die Übersetzung von akustischen Signalen in solch abstrakte Zahlenfolgen heißt Analog-Digital-Wandlung, deren Vorgang wie folgt aussieht:

Die Umwandlung eines analogen Signals in ein digitales Signal geschieht meist in zwei Schritten. Im ersten Schritt wird das am A/D-Wandler anliegende Analog-Signal in immer gleichen Zeitabschnitten abgetastet. Die Abtast-Geschwindigkeit wird Sampling-Frequenz genannt. Eine Sampling-Frequenz von 44100 Hz bedeutet, dass das Analog-Signal 44100 Mal pro Sekunde abgetastet wird. Im zweiten Schritt, den man als Quantisierung bezeichnet, wird das analoge Signal, das beliebig viele Werte annehmen kann, so verändert, dass nur eine endliche Menge von Werten vorkommen kann. Die Anzahl der Raster-Werte nennt man auch Auflösungsgenauigkeit oder kurz Auflösung (Bit). Bei der Quantisierung muss der Wandler eine Reihe von Eigenschaften erfüllen:

- wird der Analogwert größer, darf der Digitalwert nicht kleiner werden;
- wird der Analogwert kleiner, darf der Digitalwert nicht größer werden;
- bleibt der Analogwert zeitlich konstant, so darf sich auch der digitale Wert nicht ändern.

Der Analog-Digital-Wandler setzte den Startschuss für das Medium, das fortan sowohl die Compact Cassette, also auch die Vinylplatte ablösen sollte: die Compact Disc (kurz: CD). Die Informationen der CD, das sogenannte "Programm", sind auf einer spiralförmig nach

<sup>&</sup>lt;sup>46</sup> Vgl. http://www.patent-de.com/20091022/DE112007003063T5.html (Stand 13.07.2014)

<sup>&</sup>lt;sup>47</sup> Das Binärsystem ist ein Zahlensystem, das zur Darstellung von Zahlen nur zwei verschiedene Ziffern (o und 1) benutzt.

außen verlaufenden Spur angeordnet, die Abtastung erfolgt kontaktlos über einen der Spur nachgeführten Laser. Auf der Funkausstellung 1981 in Berlin wurde die CD erstmals öffentlich vorgestellt. Im sogenannten Red Book legten die Firmen *Philips* und *Sony* gemeinsam die technischen Spezifikationen der Compact Disc fest, diese beinhalteten unter anderem Stereo-Aufnahmen mit 16-Bit Auflösung bei einer Abtastrate von 44,1 kHz. 1983 wurden die ersten CD-Player für den Heimgebrauch verkauft.

Die CD trat einen globalen Siegeszug an, wenngleich sie ihrem Vorläufer, der Compact Cassette, in einem Punkt unterlegen war. Das Hauptproblem der CD bestand damals in der fehlenden Aufzeichnungsmöglichkeit für Heimanwender auf CD-Medien. Die entsprechende Technik war noch nicht ausgereift genug, um sie auf dem Markt zu veröffentlichen. Darum arbeitete man an einer Alternativlösung. 1987 wurde als Ergebnis DAT (Digital Audio Tape) eingeführt. Die Speicherung deren Informationen erfolgt auf einem Magnetband. Qualität und Komfort sind gegenüber analogen Kassetten-Recordern aufgrund der digitalen Aufzeichnung deutlich erhöht. Mit einer Abtastrate von bis zu 96 kHz und einer Auflösung von maximal 24 Bit ist das Klangpotenzial sogar höher als das der Audio-CD.

Während der DAT-Recorder für den Gebrauch im Massenmarkt auf wenig Interesse stieß, fand das System seine Nische im Bereich der Tonstudios. Diese nutzen das Digital Audio Tape für eine direct-to-stereo Aufnahme, welche in der Regel der finale Mix eines Multitrack-Projekts<sup>48</sup> war. Der DAT-Recorder bot aufgrund seiner digitalen Technik und des günstigeren Preises für diesen Schritt erhebliche Vorteile gegenüber der Bandmaschine.

Als Weiterentwicklung des zwei-kanaligen DAT-Recorders führte die Firma Alesis 1992 das Alesis Digital Audio Tape oder kurz ADAT ein. Das ADAT-Band verfügte über acht Spuren und war kompatibel mit dem kurz darauf veröffentlichten Audiorecorder ADAT XT 8. Der Clou dabei: Über ADAT-Sync-Kabel ließen sich bis zu 16 ADAT-Recorder zu einem System mit insgesamt 128 Spuren absolut phasensynchron verbinden. Als Konkurrenzprodukt brachte die Firma TASCAM im Jahr 1993 das Digital Tape Recording System (DTRS), das nach demselben Prinzip funktionierte, auf den Markt.

<sup>&</sup>lt;sup>48</sup> Vgl. Rudolf, Jr. Leonard & Leonard (2001)

Doch auch diese Technik sollte bald von einem Gerät überholt werden, das sowohl die Qualität des Digital-Audio-Tape, als auch die Handhabung der Audio-Recorder in den Schatten stellen sollte: dem Personal Computer oder kurz PC. Der PC bot die revolutionäre Möglichkeit über die auf einem Bildschirm übertragene Oberfläche mit Maus und Tastatur zu navigieren und Daten auf einer Festplatte zu speichern. Je nach Größe der Festplatte sind der Größe der Daten beziehungsweise der Länge der aufzunehmenden Audio-Signale keine Grenzen gesetzt.

Im Zuge dieser Entwicklung entstehen erste audio-basierte Programme für den Computer (in der Fachsprache auch Digital Audio Workstation oder kurz DAW genannt), allen voran die Software Sound Tools, welche die Firma Digidrums, im Jahr 1989 für den Apple Macintosh entwickelte. Mit dieser DAW war es möglich, Mono- oder Stereo-Spuren aufzunehmen, zu editieren und wiederzugeben.

Darüberhinaus bot die Software noch mehr Möglichkeiten: Durch den Einsatz des digitalen Signalprozessors (engl. digital signal processor, **DSP**) konnten die Audio-Signale in der Software in Echtzeit mit Effekten wie Equalizer, Echo oder Hall belegt werden. Einzelne Spuren ließen sich spielend leicht im integrierten Mischpult in Balance bringen und die Limitierung der Anzahl der Aufnahmespuren gehörte der Geschichte an. Zusätzlich musste man sich im Bezug auf die Schnitttechnik fortan nicht mehr nur auf den Höreindruck verlassen: Durch die visuelle Abbildung der exakten Wellenform war sofort erkennbar, an welcher Stelle der jeweilige Schnitt, der von nun an per einfachem Mausklick ausgeführt wird, gesetzt werden kann.

"We were showing people, that you could stream audio endlessly off a hard drive and do very fast nondestructive edits to it. It was a huge hit. {...} We had successfully integrated the entire recording studio into the box."<sup>49</sup>

Im Laufe der Jahre entwickelten sich die Funktionen der DAW's - wie auch die Leistung der Computer - stetig weiter. Bereits ein Jahr nach der Einführung von "Sound Tools" präsentiere die Firma Digidrums, die sich inzwischen zu Digidesign umbenannt hatte, die Nachfolge-Software "Pro Tools". Kurz darauf folgte die Firma Emagic mit ihrer DAW

24

<sup>&</sup>lt;sup>49</sup> Zit. Evan Brooks (Mitgründer der Firma Digidrums) in Milner (2010)

"Logic". Aufnahme, Bearbeitung und Reproduktion von Musik fanden jetzt in einer völlig anderen Welt statt als bisher.

Es war die Welt des Computers: Nicht mehr die Nadel und das Magnetband, sondern der integrierte Schaltkreis, der Chip im Herzen des Computers machte die Musik. Die digitale Revolution war perfekt.

#### 2. Die moderne Vocal Produktion

Wenn man sich die Anfänge der Tonaufzeichnung vor Augen hält, wird deutlich wie fortgeschritten und ausgereift die heutige Produktionstechnik ist. Wurden für Aufnahmen damals teure und sensible Wachswalzen mit einer maximalen Spielzeit von 4 Minuten verwendet, drückt man heute nur auf eine Taste an seinem Computer und die Aufnahme startet auf Anhieb und ohne Einbußen der Aufnahmequalität. Gab es lange Zeit nicht einmal die Vorstellung, dass man beispielsweise eine falsch gespielte Note eines Instruments einfach von einem anderen Take ausschneiden und einfügen könnte, geht dies heute ebenfalls spielend leicht per Mausklick.

Doch welche Möglichkeiten bietet die heutige Zeit tatsächlich? Mit dieser Frage beschäftigt sich das folgende Kapitel und bezieht sich dabei auf die Vocal Produktion. Natürlich bestehen viele Parallelen zwischen der Vocal Produktion und der Produktion mit anderen Instrumenten. Dennoch ist die Stimme das sensibelste, vielseitigste und schwierigste Instrument, anhand dessen sich ideal zeigen lässt, was uns die moderne Technik an Vorteilen - und vielleicht auch an Nachteilen - bieten kann.

# 2.1 Aufnahme

Zunächst ist es wichtig zu erwähnen, dass die Rolle des Produzenten und die des Engineers sich im Laufe der Geschichte gewaltig geändert hat. Zwar sind die Aufgaben des Produzenten (die musikalische Vision zu verwirklichen) und die des Engineers (für die technischen Aspekte für diese Verwirklichung verantwortlich zu sein) dieselben geblieben, dennoch sind diese Kompetenzen heute in vielen Fällen in einer Person vereint.

Die heutigen Aufnahme-Sessions von Vocals unterscheiden sich immens von denen der vergangenen Jahrzehnte. Durch die Digitalisierung der Tonstudios und der Ablösung der Bandmaschine durch die digitale Aufnahme am Computer eröffnet sich beispielsweise die Möglichkeit, den Sänger per Mausklick problemlos in verschiedene Teile des Songs herein

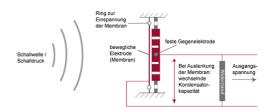
zu navigieren<sup>50</sup>. So muss der Song nicht mehr von vorne bis hinten, ähnlich wie in einer Bühnensituation, durchgesungen werden. Nun ist es also möglich, gezielt einzelne Sätze oder Worte an beliebigen Stellen mit beliebig vielen Wiederholungen aufzunehmen, den besten Take auszuwählen und die einzeln aufgenommene Stelle schließlich einzufügen. Stellt man sich diese Situation der Aufnahme-Session einmal bildlich vor, wird deutlich, welchen Aufwand der Engineer heutzutage während einer solchen Session bewerkstelligen muss. Die Flexibilität während der Aufnahmesession hat also in den vergangenen Jahrzehnten eine immer größer werdende Bedeutung erfahren.

Den Grundstein für die qualitativ hochwertige Vocal Aufnahme setzt allerdings zunächst die ausgewählte Signalkette.

#### Mikrofon

Die heutige Auswahlmöglichkeit an Mikrofonen ist nahezu unbegrenzt, wobei sich die meisten Mikrofone in der Regel in drei Typen einordnen lassen: <u>Kondensatormikrofon</u>, <u>dynamisches Mikrofon</u> oder <u>Bändchenmikrofon</u>.

Den häufigsten Einsatz bei Vocal Aufnahmen hat das <u>Kondensatormikrofon</u>. Dieses Mikrofon nutzt die durch Abstandsänderungen zwischen der (beweglichen) Membran und der Gegenelektrode hervorgerufene Kapazitätsänderungen, um je nach Ausführung Schalldruck oder Schallschnelle in ein elektrisches Signal umzuwandeln. Sie sind



**Abb. 18** Prinzip des Kondensatormikrofons als Druckgradientenempfänger

rauscharm, impulstreu, sehr empfindlich und liefern in ihrem Übertragungsverhalten einen weitestgehend linearen Frequenzgang.

<sup>&</sup>lt;sup>50</sup> diese Methode nennt man in der Fachsprache 'Punch In'. Es gab diese Methode bereits in Zeiten der Bandmaschine, dennoch kann man in der heutigen Arbeit mit DAW's die Funktion des Punch-In deutlich schneller und flexibler verwenden.

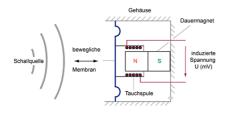


Abb. 19: Prinzip des dynamischen Mikrofons (hier ein Druckempfänger)

Ein <u>dynamisches Mikrofon</u> ist ein Mikrofon, das Schalldruckimpulse nach dem Prinzip der elektromagnetischen Induktion in äquivalente elektrische Spannungsimpulse wandelt. Sie sind mechanisch robust, relativ preiswert und benötigen keine externe Spannungsversorgung. Dynamische Mikrofone sind weniger empfindlich, das heißt sie liefern bei gleichem

Schalldruck weniger Pegel als ein Kondensatormikrofon und neigen damit stärker zum Rauschen. Dynamische Mikrofone werden vorrangig für den Live-Einsatz auf der Bühne verwendet.

Ein <u>Bändchenmikrofon</u> ist ein elektroakustischer Wandler, der wie alle dynamischen Mikrofone nach dem Prinzip der Induktion arbeitet. Beim Bändchenmikrofon sind Wandlerprinzip und akustische Bauform eng verknüpft. Bändchenmikrofone besitzen einen im Arbeitsbereich nahezu linearen Frequenzgang; die sehr geringe Masse der Membran verleiht ihnen ein gutes Impulsverhalten, weshalb sie auch als klanglicher Geheimtipp im Studio gelten.

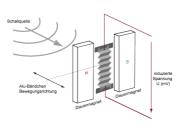


Abb. 20 Prinzip des Bändchenmikrofons

Der Klang eines Mikrofons ist unter anderem Abhängig von der Richtcharakteristik. Während sich diese bei manchen Mikrofonen umstellen lässt, sind andere mit einer festen Charakteristik ausgestattet. Die **Richtcharakteristik** beschreibt die Winkelabhängigkeit der Stärke empfangener oder gesendeter Wellen, meist bezogen auf die Empfindlichkeit bzw. Intensität in Hauptrichtung (0°). Häufig verwendete Charakteristiken sind: Niere, Superniere, Kugel oder Acht.









Kugel

Acht

Abgesehen von Mikrofonen, die auf Kugelcharakteristik eingestellt sind, verfügen alle Mikrofone über den sogenannten Nahbesprechungseffekt. Dieser Effekt ist mitunter ausschlaggebend für die Positionierung des Mikrofons und den daraus resultierenden Klang. Je näher eine Schallquelle an das Mikrofon heranrückt, desto stärker werden die tiefen Frequenzen angehoben. Wird das Mikrofon mit Lippenkontakt besungen wird eine warme Stimme erzielt. Das eignet sich für balladeske Stücke. Soll die Stimme schriller, aggressiver klingen wird das Mikrofon weiter entfernt gehalten. Aus diesem Grund gibt es nicht "die" ideale Positionierung des Mikrofons.

#### Vorverstärker

Der Vorverstärker (Preamp) ist für die Weiterverarbeitung des Mikrofonsignals verantwortlich. Während Vorverstärker oft direkt im Mischpult eingebaut sind, empfiehlt sich für die Signalkette des Mikrofons, allerdings auf ein spezialisiertes und hochwertiges externes Gerät zuzugreifen, wenngleich diese Preamps teils deutlich kostenintensiver sind. Wer noch mehr Optionen zur Klangbearbeitung sucht, findet in den sogenannten Channelstrips seinen Partner: Diese "Outboard-Geräte" verfügen neben dem Vorverstärker zusätzlich über die Funktion eines Equalizers und eines Kompressors. Eine der Hauptfragen bei der Auswahl eines Preamps liegt im Aufbau der Schaltung mit Röhren und / oder Transistoren.

#### Audio-Digital-Wandler

Für die anschließende Wandlung des analogen Signals in ein digitales Signal<sup>51</sup> benötigt man einen sogenannten Audio-Digital-Wandler. Da die Wandler, die in der Regel direkt in das

51 (für den Fall, dass man auf ein digitales Medium aufnimmt)

sogenannte Interface eingebaut sind, meist über dieselben Spezifikationen im Bezug auf die Abtastrate (kHz) und die Auflösung (Bit) verfügen, streiten Experten über die Relevanz von teureren Wandlern im Vergleich zu günstigeren Varianten.

Zusammenfassend lässt sich also sagen, dass die entscheidende Grundvoraussetzung für die Aufnahme von Vocals das Zusammenspiel aus Mikrofon, Preamp und dem Wandler ist. Auch wenn bei der Nachbearbeitung von Vocals in die Klangeigenschaften eingegriffen werden kann, ist es wichtig mit einem Signal zu arbeiten, das die optimalen technischen Voraussetzungen besitzt.

## 2.2 Vocal Comping

Bevor es später in den Bereich des Mixings geht, wird in dem Hauptteil dieser Arbeit nun eine Methode vorgestellt, mit der es möglich ist, in die Gesangsperformance des Sängers nachträglich gezielt einzugreifen. Spricht man von diesem Prinzip, ist die Rede vom sogenannten Vocal Comping.

#### **2.2.1 Prinzip**

Der Begriff *Comping* leitet sich ab von dem englischen Wort *Compositing* (Zusammensetzung). Es geht also darum, etwas zusammenzusetzen, beziehungsweise etwas zu verbinden. Beim Vocal Comping werden mehrere Aufnahmetakes zu einem einzigen verbunden.

Der Vorteil dieses Systems ist einfach erklärt: Man kombiniert die besten Stellen, die die jeweiligen Aufnahme-Takes hervorbringen und erhält somit einen "Best-Of Take". Auch wenn dieser zusammengeschnittene Take in dieser Art und Weise niemals von dem Künstler

gesungen wurde, wird der Hörer ihn nicht als zusammengesetzten, sondern als einen potenziell stimmigen und perfekt gesungenen One-Take<sup>52</sup> wahrnehmen.

Wie das geschieht zeigen folgende Beispiele:

1) Der Sänger singt einen Song drei mal vom Anfang bis zum Ende ein. Der letzte Take wird von den Beteiligten einstimmig favorisiert. Allerdings fällt auf, dass dieser in der zweiten Strophe an einigen Stellen schwach ist. Dafür ist die zweite Strophe in Take 1 überragend. Das Zusammenfügen geschieht durch zwei Mausklicks.

Ein simpler, jedoch sehr effektiver Comp ist also geschaffen.

2) Der Künstler singt die erste Strophe eines Songs fünf mal ein. Take 4 wird als Favorit gehändigt, allerdings hat der letzte Satz der Strophe nicht überzeugt, er ist zu wenig Energie geladen im Vergleich zum Rest, die Intonation zu unsauber, die Phrasierung zu sehr laid back<sup>53</sup>. Man hört sich also noch einmal die Enden aller Takes durch und sofort fällt auf: Im ersten Take passt das Ende ideal.

3) Der Künstler singt die Strophe 20mal ein. Nun geht man Schritt vor. Im ersten Satz gefällt Take 7, jedoch ist das erste Wort in Take 18 gefühlvoller eingesungen und die letzte Silbe im letzten Wort des Satzes ist in Take 14 am Besten. Nun geht man zur zweiten Phrase und zur dritten und zur vierten und durch den gesamten Song.

An dieser Stelle wird es Zeit für eine kleine Denkpause. Sicher entsteht nun der Eindruck, dass durch das anscheinend wilde Herumschneiden und Zusammensetzen das Gefühl, das der Sänger in seinen zusammenhängenden Takes zum Ausdruck gebracht hat, komplett zerstört wird und man letztlich nur noch ein rein künstliches Produkt vor sich liegen hat. Dass dieser Eindruck täuscht, wird in diesem Kapitel widerlegt.

<sup>53</sup> Fachbegriff in der Musik für eine Spiel-/Singweise, die rhythmisch zu weit hinten liegt bzw. zu spät kommt.

<sup>&</sup>lt;sup>52</sup> als One-Take bezeichnet man einen Take, der ohne Unterbrechung durchgesungen ist.

# 2.2.2 Gefühlsdefinition

Doch zuvor steht eine zentrale Frage im Raum: Wann kann man sagen, dass ein Song, ein Take oder eine Phrase gefühlvoll gesungen ist? Woran lässt sich das Gefühl messen?

Eines vorweg genommen: Das Gefühl ist ein subjektiver Gemütszustand. Dennoch muss hier muss zunächst differenziert werden. Wir reden bei der Vocal-Produktion, genauer gesagt bei der Aufnahme und dem Editing des Gesangs von zwei verschiedenen Szenarien: Wie Beurteilung seitens des Produzenten, ob der Gesang gefühlvoll war kann entweder genau dann geschehen, während der Sängern / die Sängerin den Take einsingt und somit (in den meisten Fällen) auch visuell vom Produzenten wahrgenommen wird, oder die Beurteilung findet statt, nachdem der Take aufgenommen wurde. Im letzten Fall ist die gleichzeitige Betrachtung des Sängers während seiner Gesangsperformance, also die visuelle Komponente, folglich nicht mehr vorhanden.

Doch wieso diese Differenzierung? Hier begibt man sich auf das Gebiet der Wahrnehmungspsychologie. Diese beschäftigt sich mit dem Zusammenhang von optischen und akustischen Einflüssen auf unser Gehirn und ob diese beiden Einflüsse sich in Kombination gegenseitig beeinflussen können. Im Bereich der audio-visuellen Täuschung gilt die 1976 publizierte Forschung zum McGurk-Effekt<sup>54</sup> als Meilenstein in der Wahrnehmungspsychologie. Diese Forschung belegt die Integration von visuellen Eindrücken in die auditive Wahrnehmung. Nehmen wir also den Künstler, der gerade in dem Moment singt auch visuell wahr, wird die auditive Wahrnehmung eingeschränkt beziehungsweise getäuscht. Beispielsweise ist die Beurteilung Gesangsperformance bei einem Live-Konzert rapide eingeschränkt. Hört man also die Aufnahme des Gesangs eines Live-Konzerts im Nachhinein ein weiteres Mal (bestenfalls einige wiederholte Male) an, diesmal dann ohne die visuelle Komponente der Bühnen-Situation, so fallen einem Schwächen beziehungsweise Stärken in der Stimme des Künstlers viel deutlicher auf.

Ein kurzes Beispiel aus dem Alltag für den umgekehrten Fall, also den Einfluss von auditiven Eindrücken in die visuelle Wahrnehmung und die daraus resultierende Beeinflussung der Gefühle: Man stellt sich nur einmal den Film "The Green Mile" ohne

<sup>&</sup>lt;sup>54</sup> McGurk, Harry und Macdonald, John: Hearing Lips and seeing voices. In: Nature. 264, 1976, S. 746–748

<sup>55</sup> Mehr zu diesem Thema: Jasmin Reuter: Wahrnehmung und Wirkung von Musik im Film, HDM Stuttgart, 2007

Thomas Newman's packende Musik im Hintergrund vor<sup>55</sup>. Transportiert er die Emotion der Bilder genauso intensiv auch ohne die unterlegte Musik?

Nach dieser Feststellung gibt es also eine gegenseitige Beeinflussung von auditiven und visuellen Eindrücken. So besteht für die Beurteilung aus Sicht des Produzenten, ob ein Song, ein Take oder eine Phrase gefühlvoll gesungen ist unter Umständen ein gewaltiger Unterschied, ob er den Sänger während der Aufnahme betrachtet oder die Beurteilung fernab von jener visuellen Komponenten stattfindet. Redet man über eine Song- oder Album-Produktion, redet man über eine rein auditive Wahrnehmung.

Bezogen auf den Sänger heißt das also, dass man den Gesang als gefühlvoller wahrnimmt, wenn dieser während des Singens beispielsweise die Augen geschlossen hat, emotionale Bewegungen macht, der Raum abgedunkelt ist und so weiter. All diese Dinge können zwar einen Sänger dazu verleiten, tatsächlich emotionaler zu singen. Sie können genauso aber auch den Produzenten dazu verleiten, aufgrund der visuellen Einwirkung die rein auditive emotionale Wirkung falsch zu beurteilen.

Das Szenario, dass der Produzent den Emotionsgehalt des Song, des Takes oder der Phrase zum Zeitpunkt der Aufnahme unter visuellem Aspekt auf den Sänger bestimmt, ist also nicht nur überflüssig, sondern unter Umständen auch destruktiv im Bezug auf das Endprodukt CD. Sprich: Es ist aus diesen Gründen unbedingt zu empfehlen, die Beurteilung über die gesungenen Takes erst im Nachhinein und nicht direkt während der Aufnahmen zu fällen.

# 2.2.3 Praktisches Beispiel

Unter diesem Gesichtspunkt lässt sich nun der Aspekt des Gefühls, das der Sänger in einer Aufnahme transportiert, näher betrachten. Was meinen wir also konkret damit, wenn wir behaupten, etwas ist gefühlvoll gesungen? Wie schon erwähnt ist dieses Gefühl oft schwer an etwas Konkretem festzumachen, vor allem wenn wir die visuelle Komponente außer acht lassen. Hier ein Beispiel aus der Praxis.

Schauen Sie sich nun das Video compingbeispiel.mov, das sich auf dem beiliegenden Datenträger 2 <sup>56</sup> befindet, an.

Das Video 001.mov zeigt den Teil einer Vocal Session in Logic Pro. Sie sehen vier untereinander liegende Takes einer Phrase. Die Sängerin hat diese Phrase also vier mal hintereinander eingesungen. Der Text der Phrase lautet "You've made me love, now you're making me cry". Zur Vereinfachung der Unterscheidung der Takes sind diese in verschiedenen Farben markiert, Take 1 in Rot, Take 2 in Grün, Take 3 in Gelb und Take 4 in Blau. Sie hören nun die vier Takes hintereinander und anschließend einen Comp, also einen Zusammenschnitt aus diesen vier Takes. (Für den Fall, dass dieser Comp für Sie nun aufgrund der visuellen Erkennbarkeit der *Schnitte* nicht zusammenhängend beziehungsweise zusammengeschnitten klingt, liegt auf dem Datenträger die die Audio-Datei compingbeispiel.mp3<sup>57</sup> bei, die Sie am nächsten Tag noch einmal anhören können, um zu erkennen, dass Sie dieser visuelle Eindruck getäuscht hat). Nun hören wir noch einmal die einzelnen vier Takes, aus denen der Comp stammt, um einen direkteren Vergleich zu bekommen.

Warum ich mich für die jeweiligen Takes beziehungsweise Take-Abschnitte entschieden habe möchte ich nun erläutern: In der Arbeit des Vocal Comping fokussiert ich mich in der Regel Phrasen, die eine Atem-Phase andauern. Eine solche Phrase liegt in unserem Beispiel vor. Diese Phrase wird zunächst Take für Take im Ganzen angehört, bevor man zu einer detaillierten Einteilung kommt. Meine Wahl bei "You've made me" fiel auf Take 1 (rot), allerdings klingt das "You" in Take 2 (grün) noch etwas kürzer, knackiger und somit in diesem Fall überzeugender. Das anschliessende "love" hätte ich auch von Take 1 (rot) übernehmen können, vor allem weil das vorherige "made me" ebenfalls von Rot stammt. Allerdings ist das "love" in Take 4 (blau) druckvoller und sauberer zugleich. Das "love" bei Grün und Gelb kommt für mich nicht in Frage. Bei Grün hat es zwar sehr energetisch und schön authentisch begonnen, der Abgang ist aber klar zu unsauber, Gelb hingegen ist sauber, aber zeigt im Vergleich zu Blau zu wenig Charakter. Kurioserweise spricht noch etwas dagegen, das "love" von Take 1 (rot) zu nehmen, obwohl das Vorherige bereits aus diesem Take stammt: Spulen Sie an die Stelle, in der roten Take (als Ganzes) abgespielt wird und hören Sie sich die Stelle "You've made me love" einige Male an. Schließen Sie

<sup>&</sup>lt;sup>56</sup> Datenträger 2 -> Vocal Comping -> Prinzip

<sup>&</sup>lt;sup>57</sup> Datenträger 2 -> Vocal Comping -> Prinzip

dabei die Augen. Hören Sie nun auch, dass es so klingt, als wäre das "love" unsauber geschnitten und nicht homogen mit dem Vorhergingen? Hören Sie sich im direkten Vergleich noch einmal den fertigen Comp in der Datei 001.mp3 an.

Gehen wir nun weiter zu "now you making me cry", wobei das "cry" erst einmal außer Acht gelassen wird. Im Prinzip gefallen mir hier alle Takes, es gibt jedoch einen besonderen Grund, warum ich mich für Take 2 (grün) entschieden habe. Das "me" weist hier einen besonderen Charakter auf, es ist deutlich ausdrucksstärker als alle anderen "me"-Versionen, da die Stimme für einen kurzen Moment weg bricht, tiefer geht, aber dennoch nicht an Energie verliert.

Kommen wir nun zum letzten Teil der Phrase, dem Wort "Cry". Der Vorteil, den das Vocal Comping bietet, wird nirgends so deutlich wie an dieser Stelle.

Schauen Sie sich dafür das Video compingbeispiel\_endungen.mov <sup>58</sup> an.

Wir hören die Phrase "now you making me cry" mit vier verschiedenen "cry"-Endungen. Alle vier Version sind sauber, sind gefühlvoll und haben ihren eigenen Stil. Dennoch sticht in puncto Ausdrucksstärke, Authentizität und Gefühl die Kombination Grün/Gelb deutlich im Vergleich zu den anderen Versionen heraus. Achten Sie dabei auf das Ende von "cry". Demzufolge ist die Wahl für den finalen Comp auf diese Kombination gefallen.

Wenn Sie nun den fertigen Comp der Phrase mit den ursprünglichen vier zusammenhängenden Takes vergleichen, haben Sie nicht nur bereits das Prinzip des Vocal Comping verstanden, sondern auch deutlich erfahren, welchen Mehrwert das Zusammenfügen des Materials bieten kann.

-

<sup>&</sup>lt;sup>58</sup> Datenträger 2 -> Vocal Comping -> Prinzip

# 2.2.4 Grundvoraussetzungen

Vocal Comping bietet nur unter bestimmten Bedingungen die optimalen Grundvoraussetzungen. Das Prinzip lässt sich am besten anwenden, wenn folgende Kriterien erfüllt sind:

Alle aufgenommenen Takes

- haben die gleiche Lautstärke
- besitzen die gleiche Energie im Gesangs
- sind in der identischen Position im Bezug auf Sänger <> Mikrofon aufgenommen.
- sind mit der selben Hardware (Mikrofon, Preamp, AD/DA Wandler, und so weiter) in den selben Einstellungen aufgenommen.

Sind all diese Kriterien erfüllt, so lassen sich die aufgenommenen Takes vom Prinzip her beliebig kombinieren.

Sind eine oder mehrere dieser Grundvoraussetzungen nicht gegeben, kann es Abweichungen bei der Klangfarbe der verschiedenen Takes geben. Für diesen Fall gibt es die Möglichkeit einer Reparatur im Sinne einer nachträglichen Vereinheitlichung des Klangs aller Takes.

Wenn die Aufnahmekriterien nicht einheitlich vorhanden waren, gibt es folgende nachträgliche Reparaturmöglichkeiten:

Gleiche Lautstärke: Verfügt einer der aufgenommenen Takes über eine geringere oder größere Lautstärke, so lässt sich dieser Take per Anhebung beziehungsweise Absenkung des Gains anpassen.

Gleiche Energie: Verfügt einer der aufgenommenen Takes über mehr oder weniger Energie im Gesang, so gibt es hier ebenfalls die Möglichkeit der Anhebung beziehungsweise Absenkung des Gains. Dieser Versuch der Angleichung ist jedoch kritisch zu betrachten, da er nicht immer das gewünschte Ergebnis erzielt.

Gleiche Aufnahmeposition: Ist die Aufnahmeposition des Sängers, also der Abstand und Winkel zum Mikrofon, nicht bei allen Takes derselbe, so kann es schnell enorme Abweichungen sowohl im Bezug auf die Klangfarbe, also auch auch den räumlichen Anteil geben. Ersteres ist abhängig von der Charakteristik und der Intensität des Nahbesprechungseffekts des Mikrofons. In diesem Fall hilft der Versuch der Angleichung per Equalizer. Letztes ist abhängig von der räumlichen Akustik, in dem sich das Mikrofon befindet. Je weiter sich der Sänger vom Mikrofon entfernt, umso stärker kommen die räumlichen Begebenheiten zum Vorschein. Möchte man einen Take angleichen, bei dem sich der Sänger zu nah am Mikrofon befindet, so lässt sich das unter Umständen durch den Einsatz eines künstlichen Raums (Reverb-Effekt) bewerkstelligen. Befindet sich der Sänger bei einem Take zu weit vom Mikrofon entfernt, so lässt sich der räumliche Anteil durch die Absenkung der Releasezeit in einem Transienten-Designer verringern.

Gleiche Hardware: Wird für verschiedene Takes unterschiedliche Hardware, also beispielsweise verschiedene Mikrofone oder Preamps verwendet, so lassen sich die Abweichungen in der Klangfarbe durch den Einsatz von Equalizern angleichen.

Zu betonen ist jedoch, dass die nachträgliche Angleichung der Takes eine Notfall-Lösung ist, die nicht in allen Fällen das gewünschte Ergebnis erzielen kann. Deshalb ist bei der Aufnahme darauf zu achten, dass die Grundvoraussetzungen während der gesamten Aufnahme gleich bleiben.

# 2.2.5 Parallel Comping<sup>59</sup> vs. Post Comping<sup>60</sup>

Wie bereits in Punkt 4.1.2 erwähnt, gibt es bezogen auf das Comping in der Regel zwei verschiedene Arbeitsweisen von Produzenten. Im einen Fall wird die Take-Auswahl bereits während der laufenden Aufnahme-Session vorgenommen (im Folgenden *Parallel Comping* gennant). Im anderen Fall beginnt die Take-Auswahl und das Zusammenschneiden erst im

60 Begriffsdefinition durch den Autor, Jonas Mengler, Mannheim 2014

<sup>&</sup>lt;sup>59</sup> Begriffsdefinition durch den Autor, Jonas Mengler, Mannheim 2014

Anschluss der Aufnahme-Session (im Folgenden *Post Comping* genannt). Doch welche dieser Methoden ist sinnvoller?

Um es noch einmal zu verdeutlichen: Die rein auditive Wahrnehmung durch den Endverbraucher setzt auch eine rein auditive Wahrnehmung des Produzenten während der Aufnahme heraus. Die Auswahl beziehungsweise der Zusammenschnitt der Takes beim Parallel Comping, also während der laufenden Session, bietet den Vorteil, dass direkt ein Ergebnis vorliegt und unter Umständen unsaubere Stellen direkt in der Aufnahmesession ausgebessert werden können. Allerdings wird bei dieser Arbeitsweise die laufende Session häufig unterbrochen, dies könnte den Sänger / die Sängerin aus der Stimmung bringen. Zudem könnte der Sänger das Editing während der Session beunruhigen beziehungsweise ein Zweifeln an der eigenen Gesangsperformace (selbst wenn diese exzellent war) in Gang setzen.

Editiert man erst im Anschluss an die Aufnahmesession (*Post Comping*), fallen sowohl die mögliche Verunsicherung des Künstlers auf Kosten der fortlaufenden Gesangsperformace, als auch die regelmäßigen Unterbrechnung während der Aufnahmesession weg. Zudem bietet das *Post Comping* den Vorteil, dass die Take-Kombinationen um einiges genauer ausfallen können. Beispielsweise lässt sich in der 'live-Situation' schwer sagen, welchen Endung eines bestimmten Worts in welchem Take am stärksten war. Im *Post Comping* lässt sich dies hingegen spielend leicht und vor allem zuverlässig herausfinden. Außerdem ist im Falle des *Post Comping* eine rein auditive Wahrnehmung<sup>61</sup> im Vergleich zum *Parallel Comping* garantiert.

Das Editing im Anschluss an die Aufnahmesession verlängert zwar die Produktionsdauer, dennoch bietet die Art des Post Comping klare Vorteile gegenüber dem Parallel Comping.

#### 2.2.6 Schnitte

\_

<sup>61</sup> im Sinne dessen, dass der Künstler während seiner Gesangsperformance nicht mehr visuell wahrgenommen werden kann

Redet man von Vocal Comping, so spricht man von einem direkten Eingreifen in audiobasierte Aufnahmen. Das Zerschneiden und Zusammenfügen dieser Aufnahmen muss
allerdings in einer Weise geschehen, dass beim Endverbraucher nicht der Eindruck eines
Eingriffs in das Ursprungsmatieral erweckt wird. Würde der Konsument nämlich merken,
dass in der Vocal Produktion verschiedene Takes kombiniert wurden, so wäre das Ziel der
natürlichen Optimierung der Vocal Performance verfehlt und der Mehrwert des Compings
somit verflogen. Aus diesem Grund muss jeder Schnitt, jeder Eingriff, der während des
Comping Prozesses vorgenommen wird, in einer Weise "repariert" werden, dass dieser
schliesslich nicht mehr hörbar ist. In manchen Fällen gelingt diese Reparatur problemlos,
allerdings gibt es Situationen, in denen ein gewisses handwerkliches Geschick von Nöten
ist, sowie einige Situationen, in denen es schlicht nicht möglich ist, Takes an bestimmten
Stellen zu kombinieren, ohne dass der jeweilige Schnitt zu hören ist.

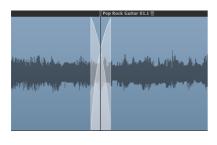


Abb. 21 Beispiel eines Crossfaders

Das zentrale Werkzeug für die Reparatur von Schnitten ist der sogenannte Crossfader. Die allgemeine Funktion eines Faders lässt sich so beschreiben: Setzt man am Ende eines Takes beziehungsweise einer Audio-Region einen Fade-Out, so blendet die Lautstärke dieser Region an der definierten Stelle aus. Mit welcher Geschwindigkeit die Lautstärke ausgeblendet wird, ist abhängig von der

gesetzten Position des Fades; je länger der Fade-Out desto langsamer die Ausblendung. Zudem lässt sich definieren, ob das Ausblenden der Lautstärke linear oder exponentiell geschehen soll. Das Prinzip des Fade-Ins ist dasselbe, allerdings bezogen auf das Einblenden eines Takes beziehungsweise einer Audio-Region.

Ein Crossfade hingegen wird eingesetzt, wenn sich zwei Regionen überlappen, also das Ende der ersten gleichzeitig mit dem Anfang der zweiten Region erklingt. Der Crossfade blendet die Lautstärke der ersten Region aus, während er die zweite einblendet (siehe Abb. 111).

Der Crossfader wird immer genau an der Stelle des Schnittes, also an der Stelle der Überlappung beider Regionen einsetzt. Dies verhindert in der Regel eine häufig eintretende Hörbarkeit der Schnitte in Form eines Knacks-Geräuschs<sup>62</sup>. Doch ein reines Setzen des Crossfades reicht nicht an allen Stellen aus. Wie beschrieben gibt es verschiedene Schwierigkeitsstufen bei der Reparatur von Schnitten. Für die Eingrenzung der Reparierbarkeit von Schnitten kommt das 4-Stufen-Modell zum Einsatz<sup>63</sup>.

# 2.2.6.1 Das 4-Stufen-Modell

# Stufe 1 - der Schnitt in Pausen

Schnitte in sogenannten Pausen, also in Stellen, an denen die Lautstärke gegen minus unendlich geht - auch als Nulldurchgang<sup>64</sup> bezeichnet - sind am leichtesten reparierbar. Ein Crossfade wird in der Regel nicht benötigt. Ist in dieser Pause allerdings ein gewisses Grundrauschen oder ein Atemgeräusch vorhanden, verhindert ein kurzer Crossfade in der Regel die Hörbarkeit des Schnittes.

#### Stufe 2 - der Schnitt auf S- und F-Lauten

Schnitte, die exakt auf S- und F-Lauten liegen, sind ebenfalls leicht reparierbar, da Schnitte an exakt diesen Stellen mit einfachem Einsatz des Crossfades in der Regel unkenntlich gemacht werden. Beispielsweise lassen sich die beiden Worte ,it seems' (,it' aus Take 1, ,seems' aus Take 2) also ohne Problem an jener Stelle verbinden, an der der "s"-Laut erklingt. Bestenfalls setzt man den Schnitt genau so, dass die erste Hälfte des "s"-Lautes von Take 1 und die zweite Hälfte von Take 2 stammt. Ein nahes Heranzoomen an die Schnittstelle innerhalb des Arrangierfensters der DAW ist hier hilfreich.

## Stufe 3 - der Schnitt an Transienten

Möchte man zwei Worte beziehungsweise Silben verbinden, bei denen die Gegebenheiten von Stufe 1 und Stufe 2 nicht vorhanden sind, so hilft ein Schnitt, der unmittelbar nach einem Transienten<sup>65</sup> liegt. Durch einen kurzen Crossfade lässt sich ein Knacksen an der

<sup>&</sup>lt;sup>62</sup> Wenn ein Schnitt an einer Stelle vollrichtet wird, an der ein Lautstärke unterschied passiert, resultiert daraus das Geräusch eines Knacksens.

<sup>63</sup> Entwickelt und formuliert vom Autor selbst, Jonas Mengler, Mannheim 2014.

<sup>&</sup>lt;sup>64</sup> Vgl. http://www.delamar.de/musik-selber-machen/audiobearbeitung/2/ (Stand: 13.07.2014)

<sup>&</sup>lt;sup>65</sup> als Transienten wird ein sehr schneller, impulshafter, elektrischer oder akustischer Einschwingvorgang bezeichnet.

Stelle des Schnitts verhindern. Ein Schnitt an einem Transienten erfordert viel Erfahrung und handwerkliche Übung.

#### Stufe 4 - der Schnitt innerhalb eines Vokals

Der Schnitt innerhalb eines Vokals ist in der Regel nicht reparierbar. Setzt man einen kurzen Crossfade, so hört man meist einen klaren Bruch an der geschnittenen Stelle des Vokals. Setzt man einen langen Crossfade, entsteht in der Regel ein sogenannter Phasing-Effekt, dabei blenden die übereinander liegenden Phasen der Singale während ihrer Überlappung verschiedene Frequenzen aus. Auch wenn der Phasing-Effekt ein gängiges Werkzeug im Bereich des Mixing darstellt, will man diesen im Bereich des Comping unter allen Umständen vermeiden, da die Natürlichkeit des gesungenen Textes unbedingt bewahrt werden muss. Das kann nur gelingen, wenn der Zuhörer den Comping Prozess nicht wahrnehmen kann. Schnitte innerhalb eines Vokals sind also zu vermeiden.

## 2.2.6.2 Praktische Beispiele für fehlerhafte Schnitte

Wie soeben erwähnt, geht in der Kunst des Vocal Compings darum, alle Eingriffe und Reparaturen absolut unhörbar vorzunehmen. Für diese Reparaturen wurde das 4-Stufen-Modell vorgestellt.

Schaut man sich die gängige Praxis anhand von weltweit erfolgreichen Songs namhafter Künstler an, so fällt auf, dass das Vocal Comping zwar im großen Maße angewandt wird, die Reparatur der Schnitte allerdings allzu oft nicht den Anforderungen an eine gute Bearbeitung entspricht.

Als Beleg dafür folgen Ausschnitte vier erfolgreicher Songs (in Form von Acappella-Versionen), die willkürlich aus einem Sortiment an hunderten Songs ausgewählt wurden. Die Beispiele liegen in Form von Videos dem Datenträger bei <sup>66</sup>. Es wird der jeweilige Song als 2-Track abgespielt, die Schnittfehler sind an den jeweiligen Stellen gelb markiert.

<sup>66</sup> Datenträger 2 -> Vocal Comping -> Schnitte

Bei Schnitt 1 ist ein Knacksen zu hören. Die Phrase lautet "You made it so sweet" (der Unterstrich auf dem Wort "made" zeigt, wo genau der Schnitt liegt). Hier wurde an einem Transienten geschnitten, jedoch auf einen Crossfader verzichtet (Stufe 3). Bei Schnitt 2 findet sich derselbe Fehler, der Schnitt liegt bei "on", direkt am Transienten ohne Einsatz eines Crossfades (Stufe 3). Schnitt 3, 4 und 5 sitzen auf der Phrase "falling from cloud nine". Im ersten dieser drei Schnitte wurde diesmal nicht an einem Transienten, sondern innerhalb des lang angehaltenen Vokals "o" des Wortes 'from' eingegriffen. Der Schnitt macht sich durch ein Knacksen bemerkbar, was darauf schließen lässt, dass hier ebenfalls auf einen Crossfade verzichtet wurde. Mit dem Einsatz des Crossfaders wäre die Hörbarkeit des Schnitts allerdings immer noch vorhanden gewesen (Stufe 4). Im zweiten dieser drei Schnitte wurde auf dem Transienten des Lauts ,c' (cloud) eingegriffen und im dritten Schnitt (Schnitt 5) am Transienten des Lautes ,i' (nine). Beide Schnitte wurden vermutlich ebenfalls ohne Einsatz eines Crossfades vorgenommen (Stufe 3). Schnitt 6 liegt in einer vermeintlichen Pause (Stufe 1), wobei die Lautstärke in diesem Fall nicht gegen minus unendlich dB geht, da ein Grundrauschen vorhanden ist. Ein einfacher Crossfade hätte hier vermutlich ausgereicht, um den Schnitt unhörbar zu machen. Schnitt 7 wäre ebenfalls leicht vermeidbar gewesen, dieser liegt zwischen den beiden Worten ,crashing from'. Wie im 4-Stufen-Modell beschrieben sind Schnitte auf F-Lauten relativ einfach durch einen Crossfade zu reparieren (Stufe 2). Allerdings müssen dabei jeweils das Ende der Region 1 und der Anfang der Region 2 exakt auf dem besagten F-Laut liegen. Hier wurde der Schnitt zu früh angesetzt. Schnitt 8 liegt innerhalb eines (langen) Vokals. Es ist möglich, dass hier ein Crossfade einsetzt wurde, da ein Knacksen nicht zu hören ist. Ist dies der Fall, wurde ein sehr kurzer Crossfade verwendet, der den Schnitt in diesem Fall aber auch nicht repariert. Wäre hingegen ein langer Fade verwendet worden, so hätte man deutlich ein Phasing-Problem gehört. In keinem der Fälle lässt sich hier also ein Schnitt reparieren (Stufe 4). Dasselbe Problem wie in Schnitt 8 findet sich auch in Schnitt 9 wieder. Schnitt 10 wurde in einer Pause vorgenommen, allerdings wäre hier durch das Grundrauschen ein kurzer Crossfade vonnöten gewesen, um diesen Schnitt zu beheben. Bei Schnitt 11 wurde unmittelbar vor anstatt unmittelbar nach dem Transienten eingegriffen (Stufe 3). Schnitt 12 hätte durch einen einfachen Crossfade bearbeitet werden können. Bei dem nächten Schnitt 13 überlappen die Worte ,nothing' der jeweiligen zwei Takes und Schnitt 14 hätte - genau wie Schnitt 12, ebenfalls durch einen einfachen Crossfade vermieden werden können.

## Beispiel 2: Kesha - Die Young

Schnitt 1 wurde an dem Transienten des Worts ,<u>b</u>eat', allerdings ohne Crossfader vorgenommen (Stufe 3). Schnitt 2 liegt in einer Pause mit einem Grundrauschen. Ein kurzer Crossfade hätte dessen Hörbarkeit verhindert (Stufe 1). Schnitt 3 liegt <u>vor</u> anstatt <u>nach</u> dem Transienten (Stufe 3). Schnitt 4 macht sich durch ein Knacksen am Transienten bemerkbar, hier wurde vermutlich auf einen Crossfader verzichtet. Derselbe Fall liegt bei Schnitt 5 vor. Schnitt 6 & 7 ist eine kopierte Stelle der vorherigen Phrasen, weshalb der Schnittfehler exakt derselbe wie bei dem vorherigen ist. Schnitt 8 liegt am Transienten des Worts ,<u>m</u>inds', das Knacksen ist vermutlich auf Grund des fehlenden Crossfades zu hören (Stufe 3). Schnitt 9 wurde kurz vor anstatt kurz nach dem Transienten des Worts ,<u>t</u>rouble' vorgenommen (Stufe 3). Schnitt 10 liegt zwar vermutlich an den richtigen Stelle des Transienten, allerdings fehlt ein Crossfader, woraus ein Knacksen resultiert (Stufe 3). Derselbe Fall liegt bei Schnitt 11 sowie Schnitt 12 vor.

# Beispiel 3: Jason Derulo - Don't Wanna Go Home

Schnitt 1 liegt in einer Pause und macht sich durch eine Knacksen bemerkbar. Da der Anfang des Takes, der unmittelbar nach dem Schnitt erklingt, inmitten eines Atemgeräuschs liegt, hätte vermutlich ein kurzer bis mittlerer Crossfade das Knacksen verhindert (Stufe 1). Schnitt 2 liegt kurz vor anstatt kurz nach dem Transienten des Worts "wanna", zudem wurde hier auf ein Crossfade verzichtet (Stufe 3). Schnitt 3 liegt inmitten des lang angehaltenen Vokals "o" ("home"). Selbst ein Crossfade hätte diesen Schnitt vermutlich nicht reparieren können (Stufe 4). Bei Schnitt 4 liegt der selbe Fall wie bei Schnitt 1 vor (Stufe 1). Schnitt 5 hätte leicht repariert werden können, wenn dieser exakt auf dem S-Laut des Worts "standing" gelegen hätte (Stufe 2). Der Schnitt wurde hier zu früh angesetzt. Schnitt 6 liegt auf dem Wort "matter" (gesprochen "madder"). Die jeweiligen Transienten des Lauts "d" (Ende Region 1, Anfang Region 2) liegen nicht übereinander, deshalb ist an exakt dieser Stelle kein Schnitt möglich, da derselbe Transient zwangsläufig in einer ungewollten schnellen Repetition erklingt.

# Beispiel 4: Madonna - Hung Up

Auch in diesem letzten Bespiel sind wieder einige unsaubere Schnitt zu hören. Schnitt 1 liegt auf dem Vokal ,e' des Worts ,every', hier wurde auf einen Crossfader verzichtet (Stufe 4). Schnitt 2 liegt am Transienten ,t' des Worts "that", hier kam auch kein Crossfader zum Einsatz, deshalb ist ein Knacksen zu hören. Schnitt 3 liegt in einer kurzen Pause unmittelbar vor einem Atemgeräusch, auch hier fehlt der Crossfader. Schnitt 4 liegt auf dem Laut ,t' am Ende des Worts ,night', ebenfalls ohne Crossfader (Stufe 3). Derselbe Schnittfehler liegt bei Schnitt 5 auf dem Wort ,tired (Stufe 3). Schnitt 6 liegt in der Pause zwischen den Silben ,tele' und ,phone'. Hier ist nicht nur ein Knacks-Geräusch, sondern auch eine zu weite Verschiebung der zweiten Silbe nach rechts zu hören (Stufe 1). Schnitt 7 liegt am Transienten ,t' des Worts ,but', ein kurzer Crossfade hätte vermutlich die Hörbarkeit dieses Schnittes vermieden (Stufe 3). Schnitt 8 liegt am Ende des langen Vokals ,a'. Selbst mit Hilfe eines Crossfades wäre der Schnitt an dieser Stelle schwer bis unmöglich zu beheben gewesen (Stufe 4). Schnitt 8 hätte hingegen leicht behoben werden können durch das Ansetzen des Schnitts unmittelbar nach dem Transienten ,t' des Worts ,that' mit Hilfe eines kurzen Crossfades (Stufe 3). Bei Schnitt 10 liegt derselbe Fall wie bei Schnitt 8 vor (Stufe 4) und Schnitt 11 liegt inmitten einer Pause beziehungsweise zwei Atemgeräuschen, hier hätte vermutlich ein mittlerer bis langer Crossfade die Hörbarkeit dieses Schnittes verhindert.

Zusammenfassend lässt sich ein bemerkenswertes Ergebnis feststellen. Es wurde deutlich, dass das Prinzip des Vocal Compings auch - oder gerade - bei großen Produktionen namhafter Künstler angewendet wird. Keines der vier Beispiele, die zufällig aus einem großen Pool an Songs ausgewählt wurden, wies eine saubere Arbeit auf.

Zur Verteidigung muss an dieser Stelle folgendes erwähnt werden: Einerseits ist zu vermuten, dass deutlich mehr Schnitte innerhalb der jeweiligen Vocal Produktionen vorgenommen wurden, die professionell repariert wurden und somit nicht hörbar sind. Zum anderen sind im Kontext des gesamten Songs (also die Acappella im Mix mit dem Playback) nur ein Bruchteil der Schnitte wahrzunehmen, die meisten werden dem "Durchschnittshörer" außerdem gar nicht auffallen. Des weiteren ist anzunehmen, dass die

Mängel bei der Schnittbearbeitung in keiner Weise Auswirkungen auf die Verkaufszahlen beziehungsweise den Erfolg des jeweiligen Songs hat.

Tatsache ist jedoch, dass im Fall der Reparatur <u>unsauber</u> und <u>unprofessionell</u> gearbeitet wurde. In Anbetracht der Tatsache, dass die Budgets für Produktionen von "Madonna" oder "Katy Perry" auf Grund ihres globalen Erfolgs überdurchschnittlich hoch ausfallen und somit ausreichend (bezahlte) Zeit für das Reparieren der Schnitt zu Verfügung stünde, bleibt also die Frage offen, warum in den jeweiligen veröffentlichten Songs letztlich so viele Schnittfehler zu hören sind.

# 2.2.7 Interview mit Ken Lewis

Im Rahmen dieser Masterarbeit führte der Autor Jonas Mengler ein Interview mit dem USamerikanischen Star-Produzenten Ken Lewis<sup>67</sup>. Der Fokus des Interviews liegt auf der Frage, welchen Mehrwert Vocal Comping für eine heutige Produktion bietet.

Interview, Jonas Mengler & Ken Lewis Mannheim, July 2014

Jonas: Hello Ken, you are a multi-platinum producer including 12 Grammy Winners and you work for artists like Kanye West, Jay-Z, Eminem, Usher, 50 Cent, John Legend, Lenny Kravitz and many more. You must be very busy, so thanks for your time.

Besides your work as a mixing and recording engineer you are an expert when it comes down to vocal comping. In your opinion: how important are vocals generally in a song compared to the playback, no matter if it's a simple programmed beat or an elaborate produced band-track?

**Ken:** Well, obviously if the artist doesn't make a connection with the listener then you have nothing, so while the music is important, the vocal is definitely the most important to get right. Comping definitely helps find the best moments and emotions of a performance.

Jonas: There are some producers who sort out takes during a running recording session (Parallel Comping) and some who do the hole recordings first and comping

<sup>&</sup>lt;sup>67</sup> Ken Lewis war bis zu diesem Zeitpunkt beteiligt an 12 Grammy Awards und 40 Nominierungen, sowie 71 Gold- und Platinauszeichnungen und 56 Nr. 1 Hits (Stand 13.07.2014).

afterwards (Post Comping). Which of these methods makes more sense to you and why?

**Ken:** I record then comp. I like to be in the moment focusing on listening to the artist as opposed to trying to make notes. However, I have a loose system during recording, I might drag a great take to a lower location (I don't use playlists) or I might color code the take file. But mostly I just record, then send the artist away and put it all together.

Jonas: How many takes do you usually record to make sure you to have enough material for your comping session?

**Ken:** I take as many takes as I need to until the artist feels to me like they are giving great performances. Once I hit that level I usually get 6 or 8 more takes and comp everything. tend to have artists sing a LOT, which gives me more work comping but often artists need some time to find their comfort zone and deliver great takes. Sometimes they just deliver great from the start and I don't need to record that many takes.

Jonas: In the article "Vocal Comping" on your website (iproducemusic.com) you say you listen to every line, every phrase, every word, every syllable on every track, and pick the very best pieces. Critics say that by cutting this much the natural feeling of the singer gets lost. How would you reply to this viewpoint?

**Ken:** Critics must not make records for a living. If I've done my job well as a producer, when I'm done with a comp, it sounds like one performance, the one best performance the artist ever gave in their life. Every breath perfect, every moment. With comping you can locate all of the magic, but you do have to be careful that each line goes with the others, or you are going to have a dysfunctional comp.

Jonas: In the text underneath your "Vocal Comping Lesson" video on youtube it says you were taught how to comp vocals by Mike Stone (who recorded Queen's "Boheman Rhapsody"). What was the most important advice he gave to you? What are the 3 most important pieces of advice about vocal comping you would give to other producers?

**Ken:** That was too long ago to remember specifics (20 years ago!!!)

1. Listen for the best feeling lines and moments. As long as pitch is in the ballpark it can be fixed.

2. Unless you have a performance with a lot of bends or vibrato, listen as you comp thru autotune. You'll find that a lot of choices you would not have picked without autotune because they were too off, once they get a bit of gentle correction, those sometimes turn into the best moments. I used to comp then correct pitch, but unless autotune isn't helping me, I now listen thru autotune and make my choices. I might still clean up more tuning later

Jonas: Thank you for your time, Ken & good luck with your upcoming projects.

//

# 2.3 Vocals im Mix

# 2.3.1 Kompressor, Reverb & Co.

Zunächst sei betont, dass im Gegensatz zum Vocal Comping, das unabhängig von jeder Musikrichtung immer nach demselben Prinzip zum Einsatz kommen kann, das Mixing der Vocals in den meisten Fällen genreabhängig ist, wenn es um die Einsatz von Effekten geht. Dennoch haben die letzten Jahrzehnte eine klare genreübergreifende Entwicklung im Bereich des Mixings durchgemacht. Bereits in den 70er Jahren erkannte das Label Motown das Problem, dass bei ihren Compilations jene Songs, die in ihrer Gesamtheit leiser gemischt waren neben den Lauteren beim Hörer untergingen. Daraus entwickelte sich in der Musikindustrie der sogenannte Loudness-War, das Streben nach einem immer höheren Lautheitspegel<sup>68</sup>. Da dem messbaren Lautstärkepegel im Gegensatz zum subjektiv hörbaren Lautheitspegel allerdings eine technische Grenze gesetzt ist, erfordert eine höhere Lautheit Tricks aus dem Bereich Psychoakustik, mit denen das subjektive Hörempfinden der Lautstärke erhöht wird. So bewirkt beispielsweise das Einschränken des Dynamikumfangs (Kompression) bei gleichbleibendem Lautstärkepegel ein subjektiv lauteres Empfinden der Musik.

<sup>&</sup>lt;sup>68</sup> Dieser ist allerdings nicht zu verwechseln mit dem Lautstärkepegel, welcher in diesem Fall der rein messbare Pegel ist. Der Lautheitspegel ist der subjektiv hörbare Pegel, der durch die Psychoakustik beeinflussbar ist.

Im Produktionsprozess des Masterings wird diese Kompression auf den gesamten Mix (2-Track im Falle eines Stereo-Mixes) angewendet. Allerdings wird die Kompression nicht nur auf den Mix als Gesamtes, sondern auch bereits auf einzelne Instrumente beziehungsweise Instrumenten-Gruppen - wie zum Beispiel den Vocals - innerhalb des Mixes eingesetzt. Der Einsatz des Kompressors ist also heutzutage in vielen Fällen ein nicht mehr wegzudenkendes Werkzeug im Bezug auf das Mixing.

Doch wie genau lässt sich die Funktionsweise eines Kompressors beschreiben?

Wie bereits erwähnt handelt es sich bei einem Kompressor um ein Tool, das zur Einschränkung des Dynamikumfangs eines Signals dient. Hierfür weist der Kompressor verschiedene Parameterwerte auf: den **Threshold**, der bestimmt, von welchem Signalpegel an der Kompressor das Signal bearbeitet, die **Ratio**, die bestimmt wie stark die



Abb. 22 Kompressor

Dynamikreduktion ausfällt, die Attack, die den Zeitpunkt des Eingreifens des Kompressor anhand des jeweiligen Signal-Transienten regelt, den Release, der die Zeit bestimmt, in der der das komprimierte Signal wieder zu seinem umkomprimierten Ursprungswert zurück findet sowie den Make-up Gain, der den gesamtem Pegel, der durch den Kompressor läuft, gleichermaßen anhebt<sup>69</sup>.

Neben dem Kompressor, dessen Einsatz für eine moderne Vocal Produktion einen wichtigen Grundstein darstellt, gibt es einige andere wichtige Werkzeuge, die für den Mix der Vocals angewendet werden.

EQ / Equalizer: mit diesem Tool lassen sich einzelne Frequenzbänder in der Lautstärke anheben beziehungsweise absenken. Man unterscheidet hierbei zwischen dem Parametischen EQ, dem Shelving EQ und dem Graphischen EQ. Die Frequenzbänder des Parametischen EQ bieten



Abb. 23 Equalizer

<sup>&</sup>lt;sup>69</sup> Letzteres ist notwendig, da der Kompressor die jeweiligen Signale in ihrem Pegel verringert und somit der Lautstärkepegel des Signals sinkt.

Einstellmöglichkeiten für den Gain, die Frequenz sowie die Flankensteilheit. Der Shelving EQ dient für eine globale Einstellung von hohen oder tiefen Frequenzen: Ab der gewählten Frequenz werden entweder alle darüber (High-Shelf) oder alle darunter liegenden (Low-Shelf) Frequenzen angehoben beziehungsweise abgesenkt. Beim grafischen EQ ist jedem Frequenzband ein eigener (Schiebe-)Regler zugeordnet. Diese Regler sind in der Regel vertikal angeordnet und bereits mit einer Flankensteilheit und einem bestimmten Frequenzbereich definiert.

Reverb / Hall: Unter dem Reverb oder Nachhall (umgangssprachlich auch Hall genannt) versteht man eine kontinuierliche Reflexion von Schallwellen in einem geschlossenen Raum oder in einem natürlich begrenzten Bereich. Im Bereich der Musikproduktion wird heutzutage in der Regel ein künstlicher Nachhall verwendet. Bei der Erzeugung eines



Abb. 24 Hall- bzw. Reverb-Effekt

künstlichen Raumeffekts wird ein Signal so verändert, dass der Zuhörer meint, das Signal sei in einer bestimmten Räumlichkeit entstanden und nicht im trockenen Tonstudio. Beim künstlichen Reverb unterscheidet man zwischen verschiedenen Typen: dem Federhall (Spring Reverb), der Hall-Platte (Plate), dem Echo Chamber, dem Faltungshall sowie dem Gated Reverb.

**Delay:** Der Delay Effekt beruht auf einer Laufzeitverzögerung des Originalsignals.

Das gängigste aller Delays ist das Tape-Delay.

Das Tape-Delay ist eine Form der

Laufzeitverzögerung, bei der die exakte Phase des

Ausgangsmaterial dupliziert und mit einem

definierbaren Zeitversatz wiedergegeben wird.



Abb. 25 Tape Delay

Zudem gibt es weiter Delay-Arten, die Produktionen zum Einsatz kommen dabei. Dazu gehört unter anderem das Stereo-Delay, das Sample-Delay oder das Reverse-Delay.

Wie bereits erwähnt, ist der Einsatz der Effekte abhängig vom Genre und Stil. Die Stimme in einer Jazz-Produktion braucht unter Umständen weniger Kompression als die in einer Pop-Produktion. Dafür kann die Stimme einer Jazz-Produktion unter Umständen wiederum mehr Hall benötigen als die eines Rappers. Eine Regel über den Einsatz der Effekte lässt sich also nicht definieren, selbst innerhalb eines Genres kann dieser stark schwanken.

Auf dem Datenträger befindet sich der Ausschnitt des Songs aus dem Kapitel 4.3.1 mit dem Text "You've made me love, now you're making me cry". Die auf dem Datenträger 2 beiliegende Datei *vocals\_miteffekten*<sup>70</sup> beinhaltet den Mix der Vocals mit Effekten, die Datei *vocals\_ohneeffekte* den Mix ohne Effekten. Für die Vocals wurden folgende Effekte verwendet: ein Equalizer, der ein wenig die hohen Frequenzen anhebt, der L3 MutliMaximer der Firma *Waves*, der als Kompressor dient und gleichzeitig eine gewisse Färbung in die Stimme bringt, ein weiterer Kompressor, der für die reine Peak Reduction zuständig ist, ein Transient Designer, der die jeweiligen Attacks minimiert und die Releases verstärkt, ein kurzer Hall mit einer Nachhallzeit von 0.8 Sekunden, ein mittlerer Hall mit einer Nachhallzeit von 1,7 Sekunden, ein Slap Delay sowie ein Stereo 8tel Delay.

# 2.3.2 Dopplungen

Ein weiteres Tool, das in einer modernen Vocal Produktion häufig angewendet wird, sind Dopplungen. Als Dopplungen bezeichnet man das Hinzufügen von einer oder mehrerer Stimmen zur Hauptstimme. Wichtig ist hierbei, dass die hinzugefügte(n) Stimme(n) exakt dasselbe singen wie die Hauptstimme. Nicht nur der Text, sondern auch jede Pause, jede Wortendung, jede Betonung und jegliche Dynamik soll in diesem Fall identisch zur Hauptstimme sein. Man erreicht eine exakte Gleichheit meist erst im Editing, indem man versucht, die Stimmen durch Schneiden und Herumschieben so gut es geht anzugleichen. Falls nun die Frage aufkommt, warum man nicht einfach die Hauptstimme kopiert, wenn die Stimmen sich sowieso gleichen sollen, gibt es eine einfache Antwort: Würde man die

<sup>&</sup>lt;sup>70</sup> Datenträger 2 -> Vocals im Mix -> Kompressor, Reverb & Co.

Stimmen kopieren, so würde kein neuer Effekt entstehen, außer dass die klingende Stimme automatisch lauter werden würde. Das erklärt sich durch die exakte Gleichung der Phase. Kopiert man also lediglich die Hauptstimme, so besitzt die Stimme für die Dopplung die exakt gleiche Phase. Legt man zwei gleiche Phasen übereinander, so summiert sich lediglich die Lautstärke. Wählt man für die Dopplung(en) jedoch einen neuen Take aus, so ist es unmöglich denselben Phasenverlauf der Hauptstimme zu erreichen. Hier kommt der Effekt der Dopplungen ins Spiel, der sich im Prinzip in drei Arten differenzieren lässt.

# Panorama Dopplungen<sup>71</sup>

Panorama Dopplungen (oder auch Stereo Dopplungen genannt) basieren auf zwei Stimmen, die zu der Hauptstimme addiert werden und wurde erst möglich durch die Erfindung der Stereophonie. Eine Stimme wird im Panorama links positioniert, die andere rechts. Wie weit das Panorama jeweils ausfällt und wie laut die Dopplungen im Verhältnis zur Hauptstimme gemischt werden, ist eine subjektive Entscheidung. Jedoch: Je stärker der Effekt der Panorama Dopplung klingen soll, desto breiter sollte das Panorama (im Extremfall ganz links beziehungsweise ganz rechts) und desto näher sollte die Annäherung der Lautstärke an die Hauptstimme (im Extremfall gleich laut mit der Hauptstimme) sein. Das Ergebnis der Panorama Dopplung ist eine Verbreiterung der Hauptstimme im Bezug auf das Stereo-Bild.

# Center Dopplungen<sup>73</sup>

Eine weitere Möglichkeit die Hauptstimme zu doppeln ist der Einsatz von nur einer weiteren Stimme, die im Panorama zentral (also genau wie die Hauptstimme) positioniert wird. Je weiter sich die Dopplung in Bezug auf die Lautstärke an die Hauptstimme annähert, desto stärker ist der Effekt. Das Ergebnis, das mit der Center Dopplung erzielt wird, wirkt deutlich kompakter als das alleinige Erklingen der Hauptstimme.

<sup>&</sup>lt;sup>71</sup> Begriffsdefinition durch den Autor, Jonas Mengler, Mannheim 2014

<sup>72</sup> Vgl. Kapitel 1.5

<sup>73</sup> Begriffsdefinition durch den Autor, Jonas Mengler, Mannheim 2014

# Stereogenerierte Center Dopplung<sup>74</sup>

Aufbauend auf der Center Dopplung gibt es eine Alternativariante, bei der diese alleinige Dopplung mit einem sogenannten Sample-Delay versetzt wird. Ein Sample-Delay ist eine Form der Laufzeitverzögerung, bei der die exakte Phase des Ausgangsmaterials dupliziert und mit einem Zeitversatz im Bereich von Millisekunden abgespielt wird. Dieser Effekt generiert die Wahrnehmung einer Panorama-Breite, beziehungsweise eines Stereo-Effektes, obwohl das Ausgangsmaterial aus einem Mono-Signal besteht.

Im Ordner "Dopplungen"<sup>75</sup>, der sich auf dem beiliegenden Datenträger 2 befindet, sind verschiedene Mixe derselben Vocal-Phrase vorhanden: Die Acappella-Version ohne Dopplungen, mit der Panorama Dopplung, mit der Center Dopplung, sowie mit der stereogenerierte Center Dopplung und zusätzlich die selben Varianten im Mix mit dem Playback.

# 2.3.3 Melodyne, Autotune & VocALign

Für die Nachbearbeitung von Vocals gibt es weitere Werkzeuge, die innerhalb der Produktions-Szene von großer Bedeutung sind.

#### Melodyne

Melodyne ist ein Tool zum Manipulieren von Tönen. Neben einer eigenständiger Software lässt sich Melodyne als Plug-In in die gängigen DAW's wie Logic, Pro-Tools & Cubase einbinden. Letzteres bietet den Vorteil, dass direkt mit Aufnahmen innerhalb der DAW gearbeitet werden kann. Melodyne dient

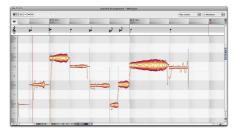


Abb. 26 Melodyne

<sup>&</sup>lt;sup>74</sup> Begriffsdefinition durch den Autor, Jonas Mengler, Mannheim 2014

<sup>&</sup>lt;sup>75</sup> Datenträger 2 -> Vocals im Mix -> Dopplungen

vorrangig der Intonationskorrektur von Gesang. Töne, die zu hoch beziehungsweise zu tief gesungen sind, können per Mausklick gerade gerückt werden. Ebenso kann die Phrasierung durch eine Kürzung, Verlängerung oder Verschiebung der Töne beziehungsweise Silben verändert werden.

#### **Autotune**

Autotune bietet so wie Melodyne die Möglichkeit der Tonhöhenkorrektur. Als Plug-In in eine DAW eingebunden, werden nach dem Einstellen der Tonart die Tonhöhen automatisch korrigiert. Dabei analysiert das Programm die Tonhöhe eines einstimmigen Audiosignals und prüft, ob und wie weit die gefundene Frequenz vom nächsten korrekten Halbton der



Abb. 27 Autotune

vorgegebenen Tonart entfernt ist, und zieht das Signal auf die korrekte Tonhöhe. Die Intensität der Korrektur lässt sich separat regulieren. In der extremen Anwendung auf eine Vocal Spur entsteht der typische "Autotune-Effekt", wie er erstmal im Song 'Believe' (1998) der Sängerin Cher zu hören war.

# VocALign

VocALign ist ein Werkzeug zum rhythmischen Angleichen von zwei Audiosignalen. In die DAW als Plug-In integriert, analysiert das Tool die Phase einer Audiodatei und gleicht diese

einer beliebigen zweiten Datei an. Sinnvoll ist dies vor allem für das Angleichen von Dopplungen. Abweichungen, die in der zweiten Datei vorhanden sind, werden durch Time-Stretching behoben. Time-Stretching beschreibt das Ändern der Wiedergabegeschwindigkeit von vorhandenem Audiomaterial, ohne dabei die Tonhöhe zu verändern.



Abb. 28 VocALign

Diese drei Tools sind hilfreich, um die Intonation, die Phrasierung sowie die rhythmische Angleichung zweier Takes für Dopplungen zu verbessern. Allerdings kann der Einsatz dieser Werkzeuge gleichzeitig eine Verfremdung des jeweiligen Audiomaterials mit sich bringen. Im Gegensatz zum Einsatz dieser Plug-Ins lässt sich die Verbesserung der Intonation, die Phrasierung sowie die rhythmische Angleichung auch auf einem anderen Wege und in einer natürlicheren Art und Weise erreichen: durch Vocal Comping. Liegt beispielsweise die Intonation einer bestimmten Note beziehungsweise Silbe zu tief oder zu hoch, so findet sich in den meisten Fällen in einem anderen Take dieselbe Silbe auf der exakteren Tonhöhe. Natürlich lassen sich falsche Töne so nicht wahllos einfügen und korrigieren, dennoch bietet dies oft die Möglichkeit mit etwas Geschick eine Note durch Ausschneiden und Einfügen von einem anderen Take zu korrigieren. Sind die Schnitte dabei reparierbar und der Kontext stimmig, so behält man gleichzeitig die Natürlichkeit des Gesangs bei, da man keine Veränderung in einer Phase vorgenommen hat sondern sich stets auf das originale Ausgangsmaterial berufen hat.

Ebenso ist dies der Fall bei der rhythmischen Angleichung zweier Audiospuren. Die minimale manuelle Änderung der Phrasierung einer Gesangsaufnahme lässt sich in der Regel durch etwas Geschick so verändern, dass der Eingriff letztlich nicht hörbar ist. Dabei wird auf jegliches Time-Stretching verzichtet und lediglich einzelne Silben an bestimmten Stellen verrückt und durch Crossfader wieder eingefügt. Während der Eingriff durch Time-Stretching in den meisten Fällen klar herauszuhören ist, kann die manuelle Änderung bis zu einem gewissen Grad unhörbar sein.

Es empfiehlt sich also, unstimmige Stellen zunächst im Zuge des Vocal Comping zu korrigieren und erst im Nachhinein bei Bedarf Werkzeuge wie Melodyne, Autotune oder VocALign zu verwenden, um dem Gesang den letzten Feinschliff zu verpassen. So ist sicher gestellt, dass die Natürlichkeit des Gesangs bestmöglich erhalten bleibt.

# 2.4 Optimaler Workflow für die moderne Vocal Produktion

Es wurde nun dargestellt, welche Möglichkeiten die moderne Vocal Produktion bietet und welche Aspekte entscheidend sind, um diese zu gestalten. Blicken wir zurück auf die Geschichte der Aufnahmetechnik und der Musikproduktion und reflektieren diese auf die heutige Situation, kristallisiert sich schnell die rasante Entwicklung und die daraus neu entstanden Möglichkeiten für die Produktion heraus.

Es wurde deutlich, dass bei der heutigen Vocal Produktion sowohl die Postproduktion (Vocal Comping & Mixing), als auch die Flexibilität während der Aufnahme-Session (Stichwort: Punch-In) eine entscheidende Rolle spielen. Diese Entwicklung wirkt sich ganz entscheidend auf die Funktionen und die Benutzung des Studioequipments aus. In Zeiten der Analogen Technik waren die Arbeitsschritte des Engineers begrenzt, während die Möglichkeiten eines Engineers heutzutage teilweise ins Unermessliche gehen. Stellen wir uns dazu eine heutige, potenzielle Vocal Recording Session eines Popsongs nun einmal bildlich vor:

Der Sänger möchte den gesamten Song zu Anfang dreimal durchsingen, um sich mit diesem vertraut zu machen. Im nächsten Schritt wird der Song unterteilt. Zuerst werden die Strophen einzeln aufgenommen, indem jede Strophe für sich mehrere Male eingesungen wird. Der letzte Satz der ersten Strophe soll einige Male separat aufgenommen werden, davor soll allerdings ein Teil vom vorletzten Take laufen, damit der Einstieg sauber er gelingt. Der Refrain wird ebenfalls noch einmal unterteilt in zwei oder drei Teile. Der erste aufgenommene Teil soll für den Einstieg in den zweiten mitlaufen und vorher mit einem Delay versehen werden.

Die Liste an möglichen Eingriffen, die in einer solchen Session vorgenommen werden können, ist nahezu unbegrenzt. Die heutigen Vocal-Sessions setzen ein völlig anderes Arbeiten im Vergleich zu dem analogen Zeitalter heraus. Demzufolge ist es unabdinglich sich für jeden Schritt, für jeden Klick und für jede Situation den optimalen Workflow anzueignen, denn es gibt für einen Künstler im Studio nichts Nervenaufreibenderes als

unnötige Wartezeit, die die Session unterbricht und somit den Sänger / die Sängerin schnell aus der Stimmung und dem Konzept bringen kann.

Der optimale Workflow lässt sich klar definieren: <u>Jede technische Ausführung eines Schritts</u>, die die Session in der jeweiligen Situation verlangt, sollte in maximaler <u>Geschwindigkeit ausgeführt und somit jede unnötige Wartezeit vermieden werden</u>. Im folgenden Kapitel wird aufgezeigt, wie der optimale Workflow in der DAW Logic Pro 9 für die Aufnahme, das Comping und das Mixing angewendet werden kann. Der erste Schritt dafür ist: die Vorbereitung.

# 2.4.1 Logic Template

Fokussiert man also das Ziel, dem Sänger während der Session jegliche nicht notwenige Wartezeit zu ersparen, müssen einige Schritte bereits erledigt sein, bevor die eigentliche Session beginnt:

- -> das Mikrofon sollte bereits grob eingepegelt sein.
- -> der Mix (bestehend aus dem 2-Track und seiner eigenen Stimme), den der Sänger auf seinem Kopfhörer hört, sollte bereits ausgewogen und seine eigene Stimme optional mit Hall versehen sein.
- -> das Playback, auf dem der Sänger / die Sängerin aufnimmt, sollte bereits in der DAW-Session anliegen, dieses sollte mit den jeweiligen Markern (Intro, Stophe, Prechorus, Chorus, C-Teil und so weiter) versehen werden.
- -> es sollten in der DAW bereits alle Spuren, die für den Verlauf der Session notwendig sein könnten, vorliegen. Dies beinhaltet im Idealfall Folgendes:

- 1 Eine Stereo-Spur für das Playback.
- 2 Eine Mono-Spur mit der Bezeichnung ,rec'. Auf dieser wird aufgenommen.
- 3 Sechs Stereo-Auxwege: ein kurzer, ein mittlerer und ein langer Hall, ein Slap-Delay sowie ein 8-tel, wahlweise ein 4-tel Stereo Delay. Diese belegen am sinnvollsten Bus 1-6.
- Vier weitere Stereo-Auxwege: einen für die Lead Vocals, einen für Dopplungen der Lead Stimme, einen für Harmonie Vocals sowie einen für Vocal Effekte. Sind mehrere Sänger an dem Song beteiligt, so werden dementsprechend mehr Aux-Wege benötigt. Alle Auxwege sollten bereits mit einer Effekt-Kette belegt sein, die EQ's und Dynamics und Send-Wege für den Hall (siehe Punkt 3) beinhaltet.
- Für jeden der vier Aux-Wege aus Punkt 4 jeweils eine Mono-"Ablagespur" mit Routing auf den jeweiligen Aux-Weg. Die jeweilige Ablagespur soll nach belieben dupliziert werden können und dient als "Parkplatz" für die Aufnahmetakes.
- Zwei Mono-Spuren jeweils mit der Bezeichnung ,c' (center) mit Routing auf den Lead Vocal Aux-Weg.
- Fine Mono-Spur mit der Bezeichnung ,l' (left) sowie eine Mono-Spur ,r' (rechts) mit Routing auf den Lead Vocal Dopplungen Auxweg. Die Spur ,l' steht im Panning links, die Spur ,r' im Panning rechts.
- Vier Mono-Spuren mit der Bezeichnung ,l' und vier Mono-Spuren mit der Bezeichnung ,r' mit dem jeweiligen Panning und in der Anordnung ,l' ,r' ,l' ,r' und so weiter, sowie drei Mono-Spuren mit der Bezeichnung ,c'. Alle Spuren werden auf den Harmonie Vocals Aux-Weg geroutet.
- Alle Spuren bis auf die sechs Aux-Wege aus Punkt 2 werden bereits versehen mit Send-Wegen zu Bus 1-6 aus Punkt 3. Diese Send-Wege lassen sich bestenfalls nach Anwählen der jeweiligen Spur durch 6 externe Controller-Fader regeln.
- Die drei Fenster der Werkzeugzuweisung (oben rechts) sollten folgende Werkzeuge beinhalten (von links nach rechts): Zeiger-Werkzeug, Scherern-Werkzeug, Marquee-Werkzeug. Ist das rechte der drei Fenster nicht sichtbar, so muss sie folgendermaßen aktiviert werden: Logic Pro -> Einstellungen -> Allgemein -> Bearbeiten -> Rechte Maustaste: Kann einem Werkzeug zugewiesen werden.
- ,,Drag" im Arrangierfenster sollte auf "X-Fade" gestellt sein, sowie "Snap" auf "Smart"

Der Low-Latency Modus sollte eingeschaltet sein (dies hängt allerdings von der Interface-Situation ab)
 Eine Master-Kette auf der bereits vorhanden Output-Spur.

Diese Voreinstellung kann in Logic Pro 9 als Template (Vorlage) unter *Ablage -> Als Vorlage sichern* gespeichert und nun stets bei dem Erstellen eines neuen Logic Projekts aufgerufen werden.

#### 2.4.2 Shortcuts

Ein ganz entschiedenes Tool für den optimalen Workflow sind Shortcuts. Shortcuts sind Programmbefehle, die sich über die Tastatur ausführen lassen.

Wie bereits erwähnt, definiert sich der Workflow durch die schnellstmögliche technische Ausführung eines Schritts, den die Session in der jeweiligen Situation verlangt. Viele der technischen Ausführungen lassen sich zwar scheinbar bequem per Mausklick erledigen. Dennoch muss die Maus für jede dieser Ausführungen eine bestimmte Strecke zurücklegen und ein bestimmtes Ziel (Button oder Menü) anvisieren. Werden dieselben Ausführungen über die Tastatur gesteuert, können diese mit etwas Übung oft um einiges schneller und präziser und somit effektiver ausgeführt werden. Für die Steigerung des Workflows ist es also höchster Relevanz, dass jeder Schritt, der potentiell durch die Maus betätigt wird, wenn möglich durch eine Taste beziehungsweise eine Tastenkombination am Keyboard ausgeführt wird.

# 2.4.3 Equipment

### 2.4.3.1 Computer und Zubehör

Neben dem Arbeiten mit einem selbst erstellen Templates und sinnvollen Verwendung von Shortcuts steht weitere eine zentrale Frage im Raum: Welches Equipment und welche Umgebung bietet die optimalen Gegebenheiten sowie den optimalen Workflow für eine Vocal Produktion?

Herzstück eines jeden Studios ist seit dem Zeitalter der Digitalisierung der Computer. Über je mehr Leistung dieser verfügt, umso schneller führt er seine Dienste aus. Ist zu wenig Leistung vorhanden, so entstehen oft unerwünschte Wartezeiten während einer Session, sei es beim Öffnen der DAW, beim Laden von Effekten, beim Erstellen einer neuen Spur, oder beim Zerschneiden eines Takes.

Die Leistung eines Computers definiert sich über mehrere Komponenten. Dabei spielt neben der CPU (Single-Core oder Muli-Core) vor allem der Arbeitsspeicher (RAM) und die Art der Festplatte (HDD oder SSD<sup>76</sup>) eine große Rolle.<sup>77</sup> Ebenso wichtig für die Annäherung an eine unverzögerte Ausführung der Befehle ist die Leistung der Grafikkarte, die im Idealfall über mehrere Bildschirmanschlüsse verfügt, denn: Je mehr Bildschirme ich an meinem Arbeitsplatz zu Verfügung habe, desto größer ist meine (digitale) Arbeitsfläche, desto mehr Fenster kann ich parallel geöffnet haben und desto weniger unnötige Zeit wird durch das Öffnen und Schließen dieser jeweiligen Fenster verbraucht.

Beispiel: An den Computer sind drei Bildschirme angeschlossen, diese sind nebeneinander angeordnet. Auf dem mittleren liegt das Arrangierfenster der DAW, auf dem rechten der Mixer und auf dem linken häufig verwendete Effekte oder Metering Plugins<sup>78</sup>. Würde man im Vergleich dazu über nur einen einzigen Bildschirm verfügen, so müssen die bestimmen Fenster immer manuell geöffnet und geschlossen werden, was logischerweise eine gewisse Zeit in Anspruch nehmen würde.

Doch nicht nur der Computer an sich, auch sein Zubehör spielt beim "Workflow" ein große Rolle. Geht es um die Wahl der richtigen Tastatur und Maus, so ist es von entscheidender Bedeutung, welche der jeweiligen Modelle zum Einsatz kommen.

<sup>&</sup>lt;sup>76</sup> Das Solid-State-Drive (kurz: SSD) bietet einen enormen Geschwindigkeitsvorteil gegen über dem Hard-Disk-Drive (kurz: HDD)

<sup>&</sup>lt;sup>77</sup> Vlg. http://www.elektronik-kompendium.de/sites/com/1309121.htm (Stand: 10.07.2014)

<sup>&</sup>lt;sup>78</sup> Metering Plugins sind zum Beispiel: Spectrum Analyser, Correlation Analyser, Level Meter

Auf der Abb. 29 sieht man verschiedene Tastaturen und Mäuse der Firma Apple. Da sich dieses Kapitel um die Arbeit mit der DAW Logic Pro 9<sup>79</sup> dreht, reden wir ausschliesslich von Mac-kompatibler Hardware.



Abb. 29 Mäuse und Tastaturen der Firma Apple

Bei den Tastaturen handelt es sich um die Modelle "Apple Keyboard Alu (Bluetooth<sup>80</sup>)" (*im Folgenden 'AKAB' genannt*), "Apple Keyboard Alu" (*im Folgenden 'AKB' genannt*) sowie das Modell "Apple Keyboard Weiss (Bluetooth)" (*im Folgenden 'AKWB' genannt*). Lenkt man seinen Blick auf den Tastenumfang der Tastaturen, so fällt auf, dass das Modell AKAB über deutlich weniger Tasten als seine Mitstreiter verfügt. Hält man sich nun die Relevanz der Shortcuts, die für den optimalen Workflow der Vocal-Produktion unabdingbar sind, vor Augen, wird klar welchen Vorteil die beiden anderen Tastaturen (AKA & AKWB) bieten. Im Klartext heißt das: Je mehr Tasten zu Verfügung stehen, desto mehr Shortcuts sind für diese jeweilige Tastatur belegbar. Hinzu kommt, dass das Modell AKAB eine direkt

<sup>&</sup>lt;sup>79</sup> die DAW Logic wurde im Jahr 2002 von Apple aufgekauft und wird seit dem ausschließlich für MAC OS Betriebssysteme entwickelt.

<sup>&</sup>lt;sup>80</sup> Bluetooth ist ein in den 1990er-Jahren durch die Bluetooth Special Interest Group entwickelter Industriestandard für die Datenübertragung zwischen Geräten über kurze Distanz per Funktechnik.

Shortcut-Belegung auf die Tasten F1-F12 verhindert, da diese Tasten erst durch die Kopplung der 'Fn'-Taste in der unteren linken Ecke, die stets gedrückt werden muss, aktiviert werden<sup>81</sup>. Das Modell AKAB erweist sich also in Relation zu den anderen beiden Modellen als ungeeignet für einen entsprechenden Workflow. Vergleicht man nun noch die beiden anderen Keyboards, so fallen zwei grundlegende Unterschiede auf. Der eine ist die Bluetooth-Funktion, die lediglich bei dem Modell AKWB vorhanden ist. Diese kann einen Vorteil gegenüber dem Kabelmodell in Sachen Flexibilität bieten, dies ist allerdings Geschmacksache des jeweiligen Nutzers. Der andere Unterschied ist allerdings deutlich spürbarer. Wirft man seinen Blick auf die Seitenansicht der jeweiligen Modelle, so fällt auf, dass das Modell AKWB über deutlich höhere Tasten verfügt. Diese bieten in Sachen Haptik einen enormen Vorteil gegenüber der flachen Variante, da Shortcuts in der Regel nur durch Fühlen und ohne direkten Blick auf die Tastatur verwendet werden. So bieten die höheren Tasten eine höhere Trefferquote im 'blinden' Verwenden.

Zusammengefasst lässt sich also sagen, dass das Modell AKWB im Vergleich zu seinen Mitstreitern die optimalen Voraussetzungen in Sachen Workflow bietet.

Die drei Mäuse der Modelle "Apple Magic Mouse Bluetooth" (im Folgenden 'AMagicMB' genannt), "Apple Mighty Mouse (Bluetooth)" (im Folgenden ,AMightyMB' genannt) und "Apple Pro Maus" (im Folgenden 'APM' genannt) weisen ebenfalls erhebliche Unterschiede in ihrer Funktionsweise auf. Die APM ist das einzige mit Kabel verbundene Modell. Im Gegensatz zu den Tastauren, die in der Regel an einem bestimmten Platz liegen bleiben, setzt die Maus eine deutlich flexiblere Handhabung im Bezug auf seine Bewegungen heraus. So weist die APM durch ihr Kabelbindung bereits einen deutlichen Nachteil gegenüber den beiden anderen Modellen vor. Hinzu kommt, dass die APM lediglich über eine einzige manuelle Taste verfügt. Die Bluetooth-Modelle AMightyMB und AMagicMB verfügen hingegen nicht nur über einen Links- und Rechtsklick, sondern auch über die Funktion des 360°-Scrollen. Dieses Scrollen ermöglicht das schnelle Navigieren des gesamten Flächeninhalts in jede Richtung innerhalb einer begrenzten Fenster-Oberfläche. Für das 360°-Scrollen verwendet die AMightyMB ein kleine Gummi-Kugel, die auf der Oberseite der Maus angebracht ist. Die Erfahrung zeigt allerdings, dass diese Kugel leicht verschmutzt und somit die kontinuierliche Funktion auf Dauer erheblich eingeschränkt ist. Die AMagicMB hingegen verwendet für das 360°-Scrollen anstatt einer

-

<sup>&</sup>lt;sup>81</sup> Werden die F-Tasten ohne die parallel gedrückte Fn-Taste gedrückt, so stehen diverse OS-bezogene Befehle wie zum Beispiel Bildschirm-Helligkeit oder Ausgangs-Lautstärke zu Verfügung.

haptischen Kugel eine eigene Touch-Oberfläche, die auf der gesamte Oberseite der Maus funktioniert. Diese Touch-Oberfläche bietet nicht nur ein zuverlässiges und kontinuierliches Scrollen in jede Richtung, sondern bietet noch weitere Möglichkeiten: Diese Maus kann erkennen, mit wie vielen Fingern in welche Richtung gestrichen wird beziehungsweise mit wie vielen Fingern in welcher Oberflächen-Region geklickt wird. Jede einzelne dieser Kombinationen (insgesamt mehr als 100 Befehle) lässt sich für die Ausführung einer Funktion innerhalb eines Programms zuweisen.

Der Vorteil der uneingeschränkten Funktion des 360°-Scrollens sowie die Möglichkeit an extra Zuweisungen durch die Touch-Oberfläche macht die Magic Mouse also zum absoluten Favoriten beim Thema Workflow.

#### 2.4.3.2 Arbeitsfläche

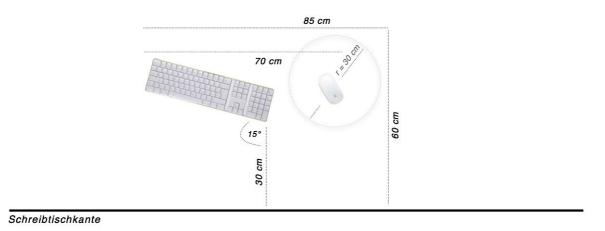


Abb. 30 Schaubild optimale Anordnung von Tastatur und Maus

Wie in Punkt 2.4.3.1 beschrieben, spielt die Kopplung aus Tastatur und Maus eine entscheidende Rolle für den Workflow am Computer. Hier ist allerdings nicht nur die Auswahl der jeweiligen Modelle von Bedeutung. Ein weiterer wichtiger Faktor für die schnelle Ausführung der technischen Befehle ist die Anordnung der Geräte sowie die Gegebenheit der Arbeitsfläche, auf der sich Maus und Tastatur befinden.

Neben einer stabilen und ebenen Oberfläche steht hier vor allem der zu Verfügung stehende Platz im Vordergrund. Die Grundposition von Tastatur und Maus wird anhand der Abbildung 30 deutlich<sup>82</sup>. Wichtig ist zunächst, dass beide Geräte mit einem Versatz von etwa 30 cm auf der Arbeitsfläche liegen, sodass der gesamte Unterarm inklusive der Ellbogen auf der Fläche aufliegen können. Der Gesamtabstand von Tastatur und Maus beziehungsweise der Ellbogen sollte etwa 70 cm betragen. Die angewinkelte Position beider Geräte (Tastatur 15° nach links, Maus 15° nach rechts) unterstützt die ergonomisch sinnvolle Ausgangslage für den Nutzer. Während die Tastatur nahezu immer an derselben Stelle positioniert bleibt, benötigt die Maus eine Kreis-Fläche von etwa 30cm Durchmesser, um bei Bedarf mit dem Mauszeiger ohne Hindernisse an jede beliebigen Stelle auf der Bildschirmoberfläche, die sich im Bestfall über mehrere Monitore erstreckt, zu gelangen. Daraus ergibt sich für die Tastatur und Maus eine optimale Arbeitsfläche von mindestens 85 x 60 cm. Ist eine stabile und ebene Fläche dieser Größe nicht gegeben, so kann es unter Umständen zu erheblichen Einschränkungen im Bezug auf den Workflow kommen.

## 2.4.3.3 Talkback-Pedal

Neben den optimalen Gegebenheiten für Tastatur und Maus gibt es ein weiteres Tool, das die Arbeitsgeschwindigkeit während einer Aufnahmesession erheblich verbessert: das Talkback-Pedal. Als "Talkback" bezeichnet man das System, mit dem derjenige im Regieraum mit dem Künstler im Aufnahmeraum auditiv kommunizieren kann. Während der Produzent im Aufnahmeraum bereits mit einem Mikrofon bestückt und so im Regieraum hörbar ist, benötigt derjenige im Regieraum ein sogenanntes "Talkback-Mic"83, sodass dieser im Aufnahmeraum über Lautsprecher oder Kopfhörer ebenfalls zu hören ist. Dieses Talkback Mikrofon ist in der Regel bereits in einem Monitorcontroller oder Mischpult am Arbeitsplatz des Produzenten oder Engineers in den Aufnahmeraum geroutet. Das Talkback-Mic ist dann aktiviert, während ein bestimmter Talkback-Knopf gedrückt ist<sup>84</sup>. In der Praxis kommt es häufig vor, dass während der Konversation zwischen den Aufnahmen der Takes parallel einige Dinge in der DAW erledigt werden können, wie zum

<sup>82</sup> die Maße beziehen sich auf einen rechtshändigen Nutzer mit durchschnittlichen Körpermaßen

<sup>83</sup> Mic steht als Abkürzung für Microphone.

<sup>&</sup>lt;sup>84</sup> Beispielsweise ist dies der Fall bei einem der meistverkauften Monitor-Controller "Big Knob" der Firma Mackie. Es gibt aber auch Systeme, bei denen ein einmaliges, kurzes Drücken auf den Talkback-Kopf das Mikrofon aktiviert bzw. deaktiviert.

Beispiel das Erstellen neuer Spuren, das Schneiden, Verschieben, Färben und Einfügen von Takes, das Inserieren von Effekten und vieles mehr. All diese Dinge benötigen unter dem Aspekt des Workflows allerdings beide Hände. Ist man also damit beschäftigt, mit einer Hand regelmäßig den Talkback-Kopf zu betätigen, behindert dies während der Konversation in den Aufnahmeraum das Ausführen parallel notwendiger Befehle in der DAW. Werden diese notwendigen Befehle (wie beispielsweise das Erstellen einer neuen Spur für den nächsten Take) nicht parallel, sondern nacheinander ausgeführt, so nehmen sie unnötig Zeit in Anspruch. Deswegen gibt es bei einigen (nicht bei allen!) Monitor-Controllern beziehungsweise Mischpulten die Möglichkeit, diese Aktivierung und Deaktivierung des Talkback-Mic's per Sustain-Pedal mit dem Fuß zu steuern und somit beide Hände frei zu haben.

Die Regelung des Talkback-Mic's per Fußpedal bietet also eine enorme Steigerung des Workflows beim parallelen Ausführen von Befehlen während der Konversation in den Aufnahmeraum.

# 2.4.3.4 Kopfhörermixer

Ein weiteres sinnvolles Gerät, das nicht nur den Workflow, sondern auch die Bequemlichkeit während der Session erleichtert, ist der Kopfhörermixer. Auf dem Kopfhörermixer wird genau das geregelt, was der Sänger in dem Aufnahmeraum auf dem Kopfhörer hört. Hier gibt es von Studio zu Studio verschiedene Systeme, die genau diesen Mix regeln. In vielen Fällen wird der Kopfhörermix vom Engineer geregelt, das bringt allerdings einige Nachteile mit sich. Angenommen der Kopfhörermixer befindet sich gekoppelt an das Audio-Interface direkt im Computer in einer extra Mixing-Software: Möchte der Sänger etwas im Mix geändert haben, so muss dieser Mixer immer geöffnet werden. Dieses beinhaltet einige Klicks und kostet Zeit und Nerven. Zudem gibt es hier ein großes Problem: Der Sänger muss dem Engineer ansagen, wie laut das Playback im Verhältnis zur eigenen Stimme und im Verhältnis zum Hall auf der eigenen Stimme sein soll. Das Ziel ist es, dem Sänger den für ihn subjektiv optimalen Mix auf die Kopfhörer zu geben. Die Einstellung des Mixes darf nicht zulange dauern, da gerade Sänger mit wenig Studioerfahrung dazu neigen, aus diesem Grund eine nicht optimale Einstellung zu

akzeptieren. Ein perfekter Kopfhörermix ist aber eine Grundvoraussetzung für das Gelingen der Gesangsperformance.

Wie also könnte der Sänger seinen Mix einfach selbst konfigurieren? Für diesen Fall gibt es eigens entwickelte Kopfhörersysteme, die beispielsweise von der Firma Axiom entwickelt wurden. Allerdings sind diese Systeme mit einem Anschaffungs-Preis von etwa 2.000 EUR sehr teuer und bieten Funktionen, die für eine reine Vocal-Produktion überflüssig sind. So sind beispielsweise für das Aviom-System 16 selbst steuerbare Kanäle belegbar. Für die Situation einer Vocal-Aufnahme werden hingegen nur drei Kanäle benötigt: die eigene Stimme (mono), das Playback (stereo) und ein Effekt-Kanal beispielsweise für Hall (stereo). Für diese Kanäle benötigt man einzig und alleine jeweils eine Lautstärke-Regelung. Im Fall des Aviom-Systems ist für die Regelung der Lautstärke der einzelnen Kanäle nur ein Drehkopf verantwortlich, der erst durch vorheriges Klicken auf einen der 16 Kanal-Knöpfe dem jeweiligen Kanal zugewiesen wird. Dies entspricht nicht dem Prinzip einer einfachen Handhabung.

# "Simplicity is the ultimate sophistication"85

Nicht nur Leonardo Da Vinci arbeitete nach dem Motto "Einfachheit ist die höchste Stufe der Vollendung", auch Apple-Gründer Steve Jobs nahm sich dieses Motto zu seiner Firmen-Philosophie.



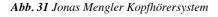




Abb. 32 Aviom Kopfhörersystem

<sup>85</sup> Zitat Leonardo da Vinci

Vlg. http://www.weyhoven.de/wie.html (Stand: 10.07.2014)

Auf der Abbildung 31 sieht man ein Gerät, das genau alle Zwecke für den selbst einstellbaren Kopfhörermix des Sänger erfüllt: einen Drehknopf für das Playback, einen für die eigene Stimme und einen für den Hall (Anmerkung: die Schalter ,Light' und ,Air' bieten keine allgemeine Relevanz für den Sänger, diese sind aufgrund der individuellen Situation der Aufnahmekabine eingebaut). In diesem System sind eigene Kopfhörerverstärker verbaut, in denen die jeweiligen Signale direkt mit Audio-Interface verbunden sind. Schaut man beide Systeme nebeneinander an, so fällt der Mehrwert des Kopfhörerverstärker in Abb. 31 schnell auf: Drei Regler, mit denen sich der Sänger auf Anhieb zurecht findet.

Ein separater, mehrkanaliger Kopfhörerverstärker, mit dem der Sänger seine Stimme, das Playback und den Hall separat und individuell selbst regeln kann, bringt also einige Vorteile mit sich.

Der optimale Workflow für eine Vocal Produktion basiert also auf einer Reihe von Komponenten. Neben der optimalen Vorbereitung der Aufnahme, die unter anderem ein DAW Template beinhaltet, das für <u>alle eventuell auftretenden Arbeitsschritte</u> bereits optimal konfiguriert ist, spielt sowohl die Leistung des Computers mit der Auswahl des richtigen Zubehörs wie Tastatur und Maus, eine stabile und ausreichende Arbeitsfläche, der bestmögliche Einsatz von Shortcuts, die Verwendung eines Talkback-Pedals sowie auch die Bedienung des Kopfhörermixers eine entscheidende Rolle. <u>Ist nur ein Teil dieser Kette nicht oder unzureichend vorhanden, so ist die Möglichkeit der Anwendung des optimalen Workflows während der Aufnahme, der Nachbearbeitung und des Mixings erheblich eingeschränkt.</u>

# 3. Fazit und Ausblick in die Zukunft

Zusammenfassend lassen sich interessante Erkenntnisse und Rückschlüsse ziehen. In dieser Arbeit habe ich aufgeführt, welchen Mehrwert die Arbeit des Vocal Compings mit sich bringt. Wie im Vorwort beschrieben, existiert bis zu diesem Zeitpunkt keine wissenschaftliche Arbeit, in der das Thema Vocal Comping im Kern behandelt wird, obwohl das Prinzip von führenden Produzenten weltweit zum Einsatz kommt. Dabei ging es zunächst nicht nur um die Erläuterung des Prinzips, sondern auch um dessen Umsetzung im optimalen Workflow. Um die Möglichkeit der modernen Schnittbearbeitung von Gesangsaufnahmen sowohl zu definieren, als auch zu differenzieren wurde im Rahmen dieser Masterarbeit das 4-Schnitt-Modell entwickelt und formuliert. Zudem wurde der Zeitpunkt, in den das Vocal Comping innerhalb einer Produktion zum Einsatz kommt mit den Begriffen "Parallel Comping" und "Post Comping" erstmals klar unterschieden und erklärt. Des weiteren habe ich anhand von praktischen Beispielen aufgeführt, dass im Bereich der Schnittbearbeitung allen vier zufällig ausgewählten Songs mit weltweitem Erfolgsstatus eine unsaubere Arbeitsweise zu Grunde liegt. Dies könnte sich unter anderem auf einen unausgereiften Workflow zurück schliessen lassen, da in diesem Fall die Schnittbearbeitung beziehungsweise die Reparatur der Schnitte ein vielfaches der Zeit im Vergleich zur Arbeit im optimalen Workflow benötigt wird.

Für die Optimierung des Workflows, die nicht nur das Vocal Comping, sondern vor allem auch die Aufnahmesession betrifft, habe ich eine Reihe von Werkzeugen und Methoden aufgeführt und durch Argumente dargelegt, dass nur die Kombination aller Werkzeuge und Methoden zum gewünschten Schnelligkeits-Ergebnis führen können.

Um nicht nur den Mehrwert von Vocal Comping, sondern auch den Mehrwert, den die gesamte heutige Produktionslandschaft zu bieten hat, deutlich zu machen, habe ich die Geschichte der Tontechnik, die beim frühen Aufnahme- und Wiedergabeverfahren des Phonautographen begann, aufgeführt.

Es ist schon beeindruckend, mit welchen Produktionsmöglichkeiten wir heute verwöhnt sind. Betrachte ich die Geschichte der Tonaufzeichnung inklusive der allerersten getätigten Aufnahme von 1860, ist es für mich wirklich erstaunlich wie weit eine Technologie in einer

verhältnismäßig so kurzen Zeit voranschreiten kann. Als Édouard-Léon Scott de Martinville im Jahr 1857 mit der Erfindung des Phonautographen den Grundstein für die Tonaufzeichnung legte, hätte dieser sich wohl kaum erträumen lassen, dass wir heute unsere Aufnahmen am Laptop tätigen, sie anschliessend spielend leicht bearbeiten, um sie direkt im Anschluss per Mausklick der gesamten Welt zu Verfügung zu stellen oder auf unser Smartphone zu übertragen.

Auch wenn der aus der damaligen Sicht undenkbare Weg vom ersten Phonautographen zum digitalen Personal Computer ein aus heutiger Sicht absolut logischer und nachvollziehbarer ist, bleibt die Vorstellung etwas Heutiges unmöglich zu Scheinendes real werden zu lassen weiterhin nur eine Illusion. Niemand kann wissen, ob wir Produzenten und Musikmachende in 100 Jahren noch immer vor einem Bildschirm sitzen und ob vor uns die Tastatur und Maus in einem optimalen Winkel von 15° positioniert ist. Niemand kann wissen, wie wir in 100 Jahren Musik konsumieren werden und vor allem wie sich die Definition des Begriffs Musik in dieser Zeit wandelt.

Fokussiere ich meine Blick zurück auf die Gegenwart und reflektiere diese auf die vergangene Entwicklung der Tontechnik, bin ich dennoch der Meinung, dass wir für die Arbeitsschritte, die von uns Produzenten heute verlangt werden - und damit meine ich beispielsweise das Vocal Comping im Bezug auf den Workflow - das Optimum an Möglichkeiten erreicht haben. An dieser Stelle möchte ich ein kurzes Beispiel nennen. Ich bin in dieser Arbeit unter anderem auf die Auswahl der richtigen Maus eingegangen und habe die Magic Mouse als das absolute Optimum betitelt. Die Magic Mouse wurde im Jahr 2009 eingeführt, zu diesem Zeitpunkt habe ich bereits seit Jahren darauf gewartet, dass eine Maus mit der Funktion des 360°-Scrollens per Touch-Funktion entwickelt wird. Ich würde schätzen, dass ich seit der Nutzung dieser Maus meinen Workflow um etwa 30% gesteigert habe. Bevor diese Maus auf dem Markt eingeführt wurde und ich bis zu diesem Zeitpunkt auf die 360° Funktion verzichten musste, war die Maus das größte Manko in meiner "Workflow-Kette". Heute sieht das anders aus. Sowohl die Maus, also auch die Tastatur und die Auswahl der konfigurierbaren Shortcuts bieten mir heute das Optimum an Workflow. Demzufolge kann ich mir zu diesem Zeitpunkt keine zukünftige technologische Entwicklung vorstellen, die - im Bezug auf das Vocal Comping - dieses Optimum überbieten könnte. Die in vielen Technologiebereichen revolutionäre Touch-Funktion wird dabei keine Rollen spielen, da unsere Finger niemals die punktierte Genauigkeit einer Spitze des Mauszeigers überbieten können, die Steuerung per Gestik wird dieses erst recht nicht erreichen. Natürlich werden im Bereich der Musikproduktion in Zukunft auch Gesten und Touch-Funktionen eine wichtige Rolle spielen. Beziehe ich mich jedoch spezifisch auf das Vocal Comping - und ich prognostiziere, dass dieses Prinzip selbst in 100 Jahren noch in dieser Art und Weise angewendet werden wird - dann sage ich voraus, dass wir heute im Jahr 2014 für diesen Bereich nahezu das technologische Optimum erreicht haben. Lediglich im Bereich der Computer-Leistung beziehungsweise der DAW-Leistung - insbesondere bei großen Projekten mit vielen Spuren und Schnitten - besteht meiner Meinung nach noch ein Potential, das zur Verbesserung des Arbeitsflusses beitragen kann.

Die Zunkuft im Bereich des Mixings wird meiner Meinung nach eine Reihe von Vereinfachungen bieten. Die Firma Waves hat mit der Einführung ihrer Signature-PlugIn-Effekten in Form eines einzelnen PlugIns, das eine Signalkette von verschiedenen Effekten bereits integriert hat, meiner Meinung nach einen Meilenstein gesetzt. Meine Erfahrung zeigt, dass die Belegung beispielsweise eines Gitarrensignals mit einem der Signature-Gitarren-PlugIns ein gleichwertigeres, wenn nicht sogar besseres, aber vor allem schnelleres Ergebnis erzeugt, wie eine eigene, "traditionelle" Zusammenstellung einer Effektkette (Beispiel: EQ, Compressor, Chorus, Delay, Hall). Ebenso erzielen die Signature-Vocal-PlugIns der Firma Waves qualitativ hochwertige Mixergebnisse in Sekunden schnelle. Da diese Art von vorgefertigten Effekt-Ketten, die nach wie vor unter dem Motto "Simplicity is the ultimate sophistication" durch wenige auserwählte Regler modifiziert werden können, stetig weiter entwickelt werden, denke ich, dass in Zukunft gerade im Bereich des Mixings viele Prozesse vereinfach werden, ohne das die Qualität des Endergebnisses darunter leidet (sie wird sich im Gegenteil dazu sogar verbessern).

Die Zukunft der Musikproduktion kann niemand vorhersagen, die Gegenwart lässt sich jedoch in den meisten Fällen wissenschaftlich analysieren, wenngleich hierbei gerade in Bereichen wie dem Vocal Comping ein massiver Nachholbedarf besteht. Ich hoffe ich konnte mit dieser Arbeit einen Teil dazu beitragen, diese wissenschaftliche Lücke zu schliessen.

# **Bonusmaterial**

Auf dem Datenträger 1 befindet sich zusätzliches Bonumaterial; zum einen der Song So Stupid, der Bereits im Rahmen der Erklärung des Vocal Comping Prinzips sowie den Mix Beispielen bekannt ist und zum anderen eine Comping Session in Form eines Videos.

# Abbildungsverzeichnis

Abb. 01 Édouard-Léon Scott de Martinville, Phonautograph, 1857	3
Abb. 02 Thomas Alva Edison mit seinem Phonographen	5
Abb. 03 Das Grammphon	6
Abb. 04 "His Masers Voice" Nipper	7
Abb. 05 Enrico Caruso während einer Aufnahme, 1902, Selbstkarikatur	8
Abb. 06 Aufnahme der Berliner Philharmoniker unter Arthut Nikisch, 1913	10
Abb. 07 1928 Neumann CMV3 Kondensatormikrofon	12
Abb. 08 Victor Orchestra bei einer Aufnahme mit einem Tricher, 1925	13
Abb. 09 Victor Orchestra bei einer Aufnahme mit einem Mikrofon, 1925	13
Abb. 10 Magnetophon K1	14
Abb. 11 Abbey Road Regieraum, 1931	17
Abb. 12 Abbey Road Studios Aufnahmeraum, 1931	17
Abb. 13 Tonbandgeräte im Maschinenraum	17
Abb. 14 Schaubild Stereophonie	18
Abb. 15 Bandmaschine Studer A800	20
Abb. 16 Kassetten Recorder der Firma Philips Typ EL 3302	21
Abb. 17 Umwandlung eines analogen Signals in ein digitales Signal	22
Abb. 18 Prinzip des Kondensatormikrofons als Druckgradientenempfänger	27
Abb. 19: Prinzip des dynamischen Mikrofons (hier ein Druckempfänger)	28
Abb. 20 Prinzip des Bändchenmikrofons	28
Abb. 21 Beispiel eines Crossfaders	39
Abb. 22 Kompressor	48
Abb. 23 Equalizer	48
Abb. 24 Hall- bzw. Reverb-Effekt	49
Abb. 25 Tape Delay	49
Abb. 26 Melodyne	52
Abb. 27 Autotune	53
Abb. 28 VocALign	53
Abb. 29 Mäuse und Tastaturen der Firma Apple	60
Abb. 30 Schaubild optimale Anordnung von Tastatur und Maus	62
Abb. 31 Jonas Mengler Kopfhörersystem	65
Abb 32 Aviam Kanfhärgrevetam	65

# Literaturverzeichnis

**Rody Rosen:** Online-Artikel "Researchers Play Tune Recorded Before Edison" (New York Times) New York, 2008.

(www.nytimes.com/2008/03/27/arts/27soun.html?pagewanted=2&\_ r=1&ref=technology)

**Haas, Walter; Klever, Ulrich:** Die Stimme seines Herrn – Eine Geschichte der Schallplatte (Ullstein Taschenbuch Verlag) Frankfurt / Main, 1959

**Roland Gelatt:** The Fabulous Phonograph. The Story of the Gramophone from Tin Foil to High Fidelity (Cassell & Company Ltd.) London, 1956.

Marcus Felsner: Operatica (Könighausen & Neumann GmbH) Würzburg, 2008

**Beadie, Jeremy J.:** Will Pop Eat Itself! Pop Music in the Soundbite Eru. (Faber B Faber) London und Boston, 1993

**Online-Artikel** "The invention of the modern music star in a hotel bedroom in Milan hervorgeht", Autor, Ort & Erscheinungsjahr unbekannt.

(http://soundofthehound.com/2011/01/16/the-invention-of-the-modern-music-star-in-a-hotel-bedroom-in-milan/)

**Stefan Gauß:** Nadel Rille Trichter: Kulturgeschichte des Phonographen und des Grammophons in Deutschland (1900-1940) (Böhlau Verlag GmbH & Cie) Köln, 2009

**Christian Preissig:** Perfect Vocals - Gesang aufnehmen und bearbeiten (PPVMEDIEN GmbH) Bergkirchen, 2009

Digitalisierung der Phonographische Zeitschriften von 1900 bis 1938

Die Phonographische Zeitschrift erschien von 1900 bis 1938 mit einem Gesamtumfang von über 34.000 Seiten. Als "Fachblatt für die gesamte Musik- und Sprechmaschinen-Industrie" (Untertitel bis 1933) stellt sie eine einzigartige Text- und Bildquelle zur Phonographie im ersten Drittel des 20. Jahrhunderts dar. Die Digitalisierung erfolgte in Kooperation mit dem Staatlichen Institut für Musikforschung Preußischer Kulturbesitz in Berlin, der Staatsbibliothek zu Berlin – Preußischer Kulturbesitz und der Bibliothek des Deutschen Museums in München

**Frank, Merlin W.:** Im Anfang war der Ton. Ein geschichtlicher Abriß über die Entwicklung in der Studiotechnik. (In: Musik Spezial Sonderheft. Home Recording und Effekte, Nr.2, S. 16-18) Grefrath, 1985

**Lutz Wernicke:** Magisterarbeit im Fach Musikwissenschaft an der Universität Osnabrück - "Musiktechnische und musikpraktische Aspekte des Homerecording-Verfahrens", 1989

Anon: BBC Hand Book 1929 (British Broadcasting Corporation; 1st edition) 1929

Peter Martland: Since Records Began – EMI: The First 100 Years (Amadeus Press) 1997

**Kim Hartmann:** Bearbeitung und Schnitt von Klassischer Musik (Tonseminar WS 09/10 Prof. Oliver Curdt) - http://www.hdm-stuttgart.de/~curdt/Klassikschnitt\_und\_Bearbeitung.pdf (Stand 07.07.2014), Stuttgart, 2009 oder 2010

**Mark Kerins:** Beyond Dolby (Stereo): Cinema in the Digital Sound Age (Indiana University Press) Bloomington, 2010

Alan P Kefauver: The Audio Recording Handbook (A-R Editions) Madison, 2001

**Hanno Ehrler:** Der Chip macht die Musik - Möglichkeiten und Konsequenzen digitaler Technologie bei der CD-Produktion, (Online-Artikel), 1996 http://www.hanno-ehrler.de/downloads/s-der\_chip\_macht\_die\_musik\_dlf.pdf

**Thomas E. Rudolph, Vincent A. Jr. Leonard & Vincent Leonard:** Recording in the Digital World: Complete Guide to Studio Gear and Software (Hal Leonard Pub Co) Milwaukee, 2001

**Andreas Friesecke:** Die Audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker (Saur, K G; Auflage: 1) München, 2009