

---

# **BACHELORARBEIT**

---

Herr  
**Nikolas Eckhardt**

**Auto-Tune, Melodyne & Co –  
Tonhöhenkorrektur und -manipulation  
am Beispiel der menschlichen Stimme**

**2019**

Fakultät: Medien

---

## **BACHELORARBEIT**

---

# **Auto-Tune, Melodyne & Co – Tonhöhenkorrektur und -mani- pulation am Beispiel der menschlichen Stimme**

Autor/in:

**Herr/Frau Nikolas Eckhardt**

Studiengang:

**Media and Acoustical Engineering**

Seminargruppe:

**MG15wA**

Erstprüfer:

**Prof. Dr. -Ing. Michael Hösel**

Zweitprüfer:

**Dipl. -Musiker Thomas Wand**

Einreichung:

**Mittweida, 21.02.2019**

Faculty of Media

---

## **BACHELOR THESIS**

---

# **Auto-Tune, Melodyne & Co – pitch correction and manipula- tion using the example of the human voice**

author:

**Mr. Nikolas Eckhardt**

course of studies:

**Media and Acoustical Engineering**

seminar group:

**MG15wA**

first examiner:

**Prof. Dr. -Ing. Michael Hösel**

second examiner:

**Dipl. -Musiker Thomas Wand**

submission:

**Mittweida, 21.02.2019**

# Inhaltsverzeichnis

<b>Inhaltsverzeichnis .....</b>	<b>I</b>
<b>Abbildungsverzeichnis .....</b>	<b>III</b>
<b>1 Einleitung .....</b>	<b>1</b>
1.1 Hinführung zur Thematik.....	1
1.2 Fragestellung .....	1
<b>2 Geschichte.....</b>	<b>2</b>
2.1 Analoge Verfahren .....	2
2.1.1 Nadeltonverfahren .....	2
2.1.2 Magnettonverfahren.....	4
2.1.3 Manipulation des Materials .....	5
2.2 Digitale Verfahren .....	7
2.2.1 Granularsynthese.....	7
2.2.2 Phase-Vocoder .....	8
2.3 Anwendungen .....	9
<b>3 Funktionsweise und Anwendungen.....</b>	<b>11</b>
3.1 technische Grundlagen .....	11
3.2 Antares Auto-Tune.....	15
3.2.1 Auto-Tune Pro.....	16
3.2.2 Auto-Tune Live.....	18
3.2.3 Auto-Tune EFX 3 und Auto-Tune 7 .....	20
3.3 Celemony Melodyne .....	21
3.4 Logic Pro X und Ableton Live 9.....	25
<b>4 Vergleich.....</b>	<b>27</b>
4.1 Vergleich gemeinsamer Features .....	28
4.1.1 Tonhöhenkorrektur.....	28
4.1.2 Tonhöhenverschiebung .....	32
4.1.3 Simulation des „Cher-Effekts“ .....	36
4.2 Funktionsumfang und Einzelfeatures.....	39
4.2.1 Melodyne .....	39
4.2.2 Auto-Tune .....	43
4.3 Benutzeroberfläche und Handhabung .....	44

---

4.3.1	Melodyne .....	45
4.3.2	Auto-Tune .....	45
4.4	Zwischenfazit .....	46
<b>5</b>	<b>Anwendungsbereiche.....</b>	<b>48</b>
5.1	Perfekte Intonation.....	48
5.2	Zeitersparnis .....	49
5.3	„Cher-Effekt“ .....	50
<b>6</b>	<b>Kritik und Entwicklung .....</b>	<b>52</b>
6.1	Kritik .....	52
6.2	Entwicklung.....	53
<b>7</b>	<b>Fazit.....</b>	<b>55</b>
	<b>Literaturverzeichnis .....</b>	<b>IV</b>
	<b>Anhang .....</b>	<b>VIII</b>
	<b>Eigenständigkeitserklärung .....</b>	<b>XI</b>

---

## Abbildungsverzeichnis

Abbildung 1: Der erste Edison-Phonograph .....	2
Abbildung 2: Magnettonverfahren.....	4
Abbildung 3: A/D-Wandlung .....	11
Abbildung 4: Basis-Ansicht Auto Mode .....	16
Abbildung 5: Ansicht Graph Mode .....	17
Abbildung 6: Ansicht Auto-Tune Live .....	19
Abbildung 7: Ansicht Melodyne .....	22
Abbildung 8: Pitch Shifter und Pitch Correction (Logic Pro X) .....	25
Abbildung 9: polyphoner und monophoner Algorithmus .....	30
Abbildung 10: Sample-Bearbeitung Ableton Live 9 Suite .....	34
Abbildung 11: Vor und nach der Tonhöhenmodulation.....	37
Abbildung 12: Notenerkennung Melodyne .....	40
Abbildung 13: mehrstimmige Analyse Melodyne .....	41

# 1 Einleitung

Wer schon mal ein Foto bearbeitet hat oder den Unterschied zwischen vor und nach dem Color-Grading eines Videos kennt, weiß, wie sehr Computersoftware Rohmaterial verbessern und verändern kann. Diese Entwicklung findet seit Jahrzehnten auch in der Musikbranche statt. Auf einen Teil dieser Entwicklung, nämlich die Korrektur und Manipulation von Tonhöhen soll am Beispiel der menschlichen Stimme im Laufe dieser Arbeit eingegangen werden.

## 1.1 Hinführung zur Thematik

Schon zu Zeiten der analogen Musikproduktion war es normal, das Audiomaterial nach der Aufnahme zu beschneiden und zu bearbeiten. Die Aufnahme wurde also im Nachhinein künstlich verbessert und verändert. Spätestens mit Beginn der digitalen Audioproduktion wurden die Möglichkeiten der Bearbeitung immer vielseitiger und komplexer. Ein Bereich der Bearbeitung ist die Veränderung der Stimme bzw. des Gesangs. Besonders im Rahmen seines Praktikums in einem Tonstudio wurde der Verfasser mit dieser Thematik konfrontiert, als er die vier Stimmen eines schlecht intonierenden Chores per Hand in die richtige Tonhöhe bringen musste. Zudem ist es gängige Praxis, dass Gesang im Nachhinein mit verschiedenen Effekten bearbeitet wird. Dazu gehört auch die Tonhöhenkorrektur, welche das Hauptthema dieser Arbeit darstellt.

## 1.2 Fragestellung

Hauptziel dieser Arbeit ist es zu klären, inwiefern sich verschiedene digitale Programme zur Tonhöhenkorrektur und -manipulation voneinander unterscheiden. Dabei wird sowohl der Funktionsumfang als auch das Klangergebnis betrachtet. Zum Vergleich werden die beiden am weitesten verbreiteten Softwares Auto-Tune von Antares und Melodyne von Celemony aufgeführt. Wie weit reichen die Möglichkeiten der Tonhöhenkorrektur und -manipulation und welche Software ist dafür am besten geeignet?

## 2 Geschichte

Ursprünglich waren akustische Signale vergängliche Momente, die nur zum Augenblick ihres Erklingsens präsent waren. Erst durch die Entwicklung technischer Methoden zur Aufnahme eines Signals war es möglich, akustisches Material festzuhalten, zu reproduzieren und in diesem Zuge auch zu manipulieren. Im Folgenden wird zunächst auf analoge Verfahren eingegangen, um später digitale Verfahren zu beschreiben.

### 2.1 Analoge Verfahren

#### 2.1.1 Nadeltonverfahren

Weit über 100 Jahre zurück liegen die ersten Versuche, eine Maschine zur Schalllaufzeichnung und -wiedergabe zu bauen. Der Erste, dem der Bau einer solchen Maschine gelang, war der Amerikaner Thomas Alva Edison im Jahr 1877.<sup>1</sup> Zwar hatte der Franzose Charles Cros bereits drei Monate zuvor Zeichnungen und

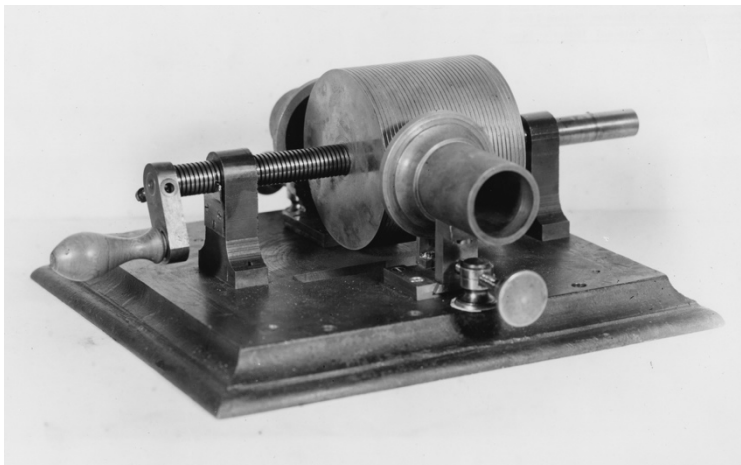


Abbildung 1: Der erste Edison-Phonograph

Beschreibungen einer ähnlichen, vom ihm „Paléophone“ genannten Maschine an der Pariser Akademie der Wissenschaft hinterlegt, ihm fehlten allerdings sowohl das Geld zur Patentanmeldung als auch jemand mit den technischen Fertig-

keiten, seine Erfindung tatsächlich zu bauen.<sup>2</sup> Der Verdienst der technischen Umsetzung gebührt also Edison und seinem Team. Mit Hilfe einer Nadel, die an eine Membran gelötet war, wurden die Schwingungen der Membran auf einen mit Paraffin überzogenen Papierstreifen übertragen. Auf umgekehrtem Weg versetzte die aufgezeichnete Schwingungslinie die Nadel und dementsprechend auch die

---

<sup>1</sup> Vgl. Große, Günter (1981): Von der Edisonwalze zur Stereoplatte. 1. Auflage, Berlin. S. 9.

<sup>2</sup> Vgl. ebd. S. 10.



Membran in Schwingung. Mit diesem Gerät, dem so genannten „Edison-Phonographen“, wurde die menschliche Stimme zum ersten Mal konserviert.<sup>3</sup>

Nur zehn Jahre später, also 1887, entwickelte der gebürtige Deutsche Emil Berliner, der zu dieser Zeit in Washington lebte, ein Verfahren, bei dem anstelle von Wachswalzen ebene Scheiben zur Schallaufzeichnung verwendet wurden. Dies realisierte er mittels einer Zinkscheibe, auf die eine dünne Wachsschicht aufgetragen war. Die Schallschwingungen wurden spiralförmig von außen nach innen eingraviert und anschließend durch Eintauchen in ein Säurebad in die Zinkplatte geätzt. Dadurch waren die Schalllinien nicht nur dauerhaft festgehalten und damit beliebig oft abspielbar, es bot sich außerdem die Möglichkeit, Kopien der Zinkplatte aus Hartgummi herzustellen. Damit war die Schallplatte geboren.<sup>4</sup> Zur Wiedergabe benötigte man ein Grammophon, das anfangs noch mit einer Handkurbel betrieben werden musste. Der kommerzielle Erfolg war allerdings anfangs gering, zumal der konkurrierende Phonograph zu dieser Zeit schon durch Uhrwerke oder kleine Elektromotoren betrieben wurde. Erst 1896 gelang es Eldridge R. Johnson, Berliners Grammophon durch ein Uhrwerk anzutreiben, das in seiner Laustärke die Musik nicht übertönte und dem so genannten „Improving Gramophone“ 1897 zum großen Durchbruch verhalf.

Durch stetige Weiterentwicklung und Verbesserung des Nadelton-Verfahrens erfreute sich die Schallplatte einer immer höheren Popularität. Mit der Entstehung des Rundfunks in den 1920er Jahren wurde immer häufiger Musik von Schallplatten auch per Radio übertragen. Wenig später revolutionierte das elektrische Aufnahmeverfahren des Rundfunks die Schallplattenindustrie. Durch Verstärker und Mikrofon anstelle der Schalltrichter wurde eine deutlich höhere Klangqualität erreicht und die Grundlage für den technischen Werdegang der Schallplatte in den nächsten Jahren gelegt.<sup>5</sup>

---

<sup>3</sup> Vgl. Große, Günter (1981): Von der Edisonwalze zur Stereoplatte. 1. Auflage, Berlin. S. 11 ff.

<sup>4</sup> Vgl. ebd. S. 28 f.

<sup>5</sup> Vgl. ebd. S. 72 f.

## 2.1.2 Magnettonverfahren

Zu der mechanischen Aufzeichnung von Schallschwingungen kam 1936 das Tonband mit magnetischer Aufzeichnung hinzu. Beim sogenannten Magnettonverfahren werden Schallsignale auf magnetisierbare Materialien aufgezeichnet, die gespeichert und wieder ausgelesen werden können.<sup>6</sup> Bekannt war dieses Prinzip bereits seit der Jahrhundertwende, allerdings war es der Wachsplatte qualitativ lange Zeit unterlegen. Der deutsche Erfinder Fritz Pfelemer verbesserte 1928 die Qualität der Aufnahme, indem er ein nichtmagnetisches Trägermaterial mit einer magnetisierbaren Schicht überzog. Es folgten weitere Verbesserungen, bis die deutschen Rundfunkanstalten das Verfahren 1938 bei sich einführten. International bekannt wurde das Magnettonverfahren allerdings erst nach 1945.<sup>7</sup>

Die magnetische Schallspeicherung, also die Aufzeichnung eines Schallsignals, erfolgt, indem der Tonträger vor dem Spalt eines so genannten Sprechkopfes vorbeigeführt wird, in dessen Wicklung Wechselstrom fließt. Der Tonträger be-

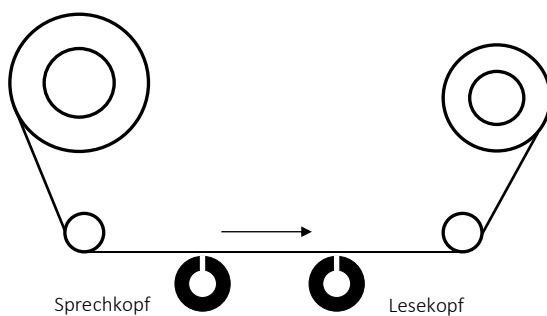


Abbildung 2: Magnettonverfahren

steht dabei aus einem magnetisierbaren Material. Das Wechselfeld, welches vor dem Spalt des Sprechkopfes entsteht, „bespielt“ den Tonträger mit einer bleibenden Magnetisierung örtlich wechselnder Größe und Richtung. Die Wiedergabe erfolgt durch das Vorbeiführen des magnetisierten Materials an einem Hör- bzw. Lesekopf, in dem eine Spannung entsprechend des magnetisierten Tonträgers induziert wird.<sup>8</sup>

Mit der Zeit wurde das Verfahren in Zusammenarbeit mit dem Rundfunk und der Industrie erheblich verbessert. Die hochfrequente Vormagnetisierung machte das Verfahren auch für hochwertige Musikaufnahmen geeignet. Durch die

<sup>6</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2008): Handbuch der Tonstudioteknik, Band 1. 7. Auflage, München. S. 376.

<sup>7</sup> Vgl. ebd.

<sup>8</sup> Vgl. ebd. S. 377.

Verwendung von breiteren Bändern wurden außerdem zunächst Stereoaufnahmen und später auch Mehrspuraufnahmen möglich.<sup>9</sup>

### 2.1.3 Manipulation des Materials

Mit der Einführung von technischen Verfahren zur Aufnahme von Audiosignalen und der damit einhergehenden Reproduzierbarkeit wurde es auch möglich, das aufgenommene Material nachträglich zu verändern. Beispielsweise konnte man durch Zerschneiden und erneutes Zusammensetzen von Magnetbändern nicht nur Fehler bei der Aufnahme beheben, sondern auch ein komplett neues Stück entstehen lassen. Ein anderer Effekt war die Manipulation der Tonhöhe durch Veränderung der Abspielgeschwindigkeit. Durch schnelleres Abspielen der Platte bzw. des Magnetbandes wurde die Tonhöhe angehoben, durch langsameres Abspielen abgesenkt. Geschwindigkeit und Tonhöhe konnten zu dieser Zeit noch nicht unabhängig voneinander verändert werden. Wollte man also die Tonhöhe eines Audiosignals nach oben oder unten korrigieren, ging dies immer mit einer Veränderung der Geschwindigkeit einher. Des Weiteren wurde bei dieser Art der Transposition auch der Klangcharakter des Materials verändert, da ja nicht nur der Grundton, sondern das gesamte Frequenzspektrum eines Klanges verschoben wird. Besonders kritisch in diesem Zusammenhang ist die Formantenverschiebung. Formanten sind von besonderen Eigenschaften des Klangerzeugers abhängige Resonanzfrequenzen im Obertonspektrum, die maßgeblich zur Gestaltung der Klangfarbe beitragen und unabhängig vom Grundton entstehen.<sup>10</sup>

In den 1950er Jahren entwickelte Axel Springer das sogenannte Tempophon. Dieses Gerät war einer der ersten Versuche, Geschwindigkeit und Tonhöhe unabhängig voneinander zu verändern. Das Gerät verfügte über vier rotierende Schreib-/Leseköpfe, die jeweils akustische „Schnipsel“ aus dem Originalmaterial entnahmen, aus denen wiederum ein kontinuierliches Signal rekonstruiert werden konnte. Indem man die Bandlaufgeschwindigkeit und die

---

<sup>9</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2008): Handbuch der Tonstudioteknik, Band 1. 7. Auflage, München. 378 f.

<sup>10</sup> Vgl. Sengpiel Audio (1994): Formanten prägen die Klangfarbe, <http://www.sengpielaudio.com/Formanten-PraegenDieKlangfarbe.pdf> (Zugriff: 11.01.2019).

Drehgeschwindigkeit der vier Köpfe veränderte und aneinander anpasste, konnten Geschwindigkeit und Tonhöhe des Materials unabhängig voneinander verändert werden. Die Qualität des ausgegebenen Signals war allerdings eher schlecht. Durch das Entnehmen und erneute Zusammensetzen der einzelnen „Schnipsel“ entstanden an den Übergängen viele Knackser und Verzerrungen.<sup>11</sup>

Ein frühes Beispiel für die praktische Anwendung von Tonhöhenverschiebung der Stimme lieferte der Komponist Paul Hindemith 1930 im Rahmen des Musikfestivals „Neue Musik Berlin“. In seinem Stück „Gesang über 4 Oktaven“ kreierte er Dopplungen und Akkorde, indem er seine selbst eingesungene Stimme in doppelter bzw. halber Geschwindigkeit abspielte.<sup>12</sup> In derselben Nacht präsentierte Hindemiths Zeitgenosse Ernst Toch seine „Gesprochene Musik“.<sup>13</sup> Dabei entstand durch immer schnelleres Abspielen eines gesprochenen Textes eine Art Melodie. Diese und ähnliche Experimente werden der so genannten Grammophonmusik zugeordnet.<sup>14</sup>

In den folgenden Jahren entstanden weitere experimentelle Musikstile, die eine Manipulation des aufgenommenen Materials als festen Bestandteil der Komposition einsetzten. Ein weiteres Beispiel ist hier die in Paris entwickelte „musique concrète“, die bevorzugt Montagen von aufgenommenen Klängen und Geräuschen für ihre Kompositionen verwendete.<sup>15</sup> Um im Bereich der Tonhöhenmanipulation der Stimme zu bleiben, ist hier insbesondere der populäre Song „The Chipmunk Song“ aus der TV-Serie „Alvin und die Chipmunks“ von 1958 zu erwähnen.<sup>16</sup> Für den hohen, kindlich verzerrten Klang der Stimmen wurde der Gesang in langsamem Tempo auf Magnetband aufgenommen und im Nachhinein schneller abgespielt. Dieser Effekt wird nach wie vor als „Chipmunk-Effekt“ bezeichnet.<sup>17</sup> Mindestens genauso bekannt und auf gleiche Weise erzeugt ist der

---

<sup>11</sup> Vgl. Peter Manning (1993): *Electronic & Computer Music*, 2. Auflage, USA, S. 82 ff.

<sup>12</sup> Vgl. Mark Katz (2010): *Capturing Sound. How technology has changed music*, London, S. 110.

<sup>13</sup> Vgl. Hörbeispiel Nr. 1.

<sup>14</sup> Vgl. Mark Katz (2010): *Capturing Sound. How technology has changed music*, London, S. 111 ff.

<sup>15</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2008): *Handbuch der Tonstudioteknik*, Band 1. 7. Auflage, München, S. 83.

<sup>16</sup> Vgl. Hörbeispiel Nr. 2.

<sup>17</sup> Vgl. Peter Manning (1993): *Electronic & Computer Music*, 2. Auflage, USA, S. 205.

Mickey-Mouse-Effekt, also die hohen Sprechstimmen der Trickfiguren in den Filmen von Walt Disney.<sup>18</sup>

## 2.2 Digitale Verfahren

Mit der Einführung und Verbreitung von digitalen Technologien zur Aufnahme von Audiosignalen machte auch die Musikbranche eine große Entwicklung durch. Die Veränderung der Tonhöhe unabhängig von der Geschwindigkeit war zwar zu analogen Zeiten schon möglich, etwa mit dem oben beschriebenen Tempophon, allerdings war dies ein sehr kompliziertes Unterfangen, unter dem auch die Qualität litt. Nachfolgend werden die Geschichte und vereinfacht das Prinzip der Granularsynthese und des Phasen-Vocoders beschrieben. Eine genauere Betrachtung von mathematisch-technischer Seite erfolgt in Kapitel 3.

### 2.2.1 Granularsynthese

Das Prinzip der Granularsynthese beruht darauf, tausende sehr kurzer Klangfragmente (so genannte „Grains“) zu einem Klang zusammenzuführen. Diese Grains, ins Deutsche sinngemäß als „Klangkörner“ zu übersetzen, fungieren also quasi als Bausteine für Soundobjekte. Mit diesem Verfahren lassen sich einerseits vorhandene Klänge zerteilen und neu zusammensetzen, andererseits können aber auch völlig neue Klänge erschaffen werden.<sup>19</sup>

Die Idee, dass ein Klang aus vielen verschiedenen „Soundpartikeln“ bestehe, existiert schon seit mehreren hundert Jahren. Der britische Physiker Dennis Gabor entwickelte daraus 1946 die Theorie, dass jeder beliebige Klang erschaffen werden könne, indem man einzelne Klangpartikel zusammensetzt.<sup>20</sup> 1960 war der griechische Komponist Iannis Xenakis der erste, der auf dem Prinzip der Granularsynthese eine kompositorische Theorie erstellte:

---

<sup>18</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2008): Handbuch der Tonstudioteknik, Band 1. 7. Auflage, München, S. 347.

<sup>19</sup> Vgl. Roads, Curtis (2004): Microsound, USA 2001. S. 87.

<sup>20</sup> Vgl. Roads, Curtis (1996): The computer music tutorial, USA. S. 168 f.

*„All sound, even continous musical variation, is conceived as an assemblage of a large number of elementary sounds adequately disposed in time. In attack, body, and decline of a complex sound, thousands of pure sounds appear in a more or less short intervall of time.“<sup>21</sup>*

Xenakis setzte diese Theorie um, indem er 1959 seine Komposition „Analogique A et B“<sup>22</sup> schuf, bei der er Klanguaufzeichnungen verschiedener Magnetbänder in winzige Teile zerschnitt und neu zusammensetzte.

Die ersten Experimente mit computerbasierter Granularsynthese machte US-amerikanische Komponist und Programmierer Curtis Roads 1974 an der University of California in San Diego. Die Berechnung dieser Kompositionen dauerte damals allerdings noch sehr lange. Erst in den späten 1980er Jahren entwicklele der kanadische Komponist Barry Truax ein Verfahren zur Berechnung granularer Klänge in Echtzeit.<sup>23</sup>

## 2.2.2 Phase-Vocoder

Das Prinzip des Phase-Vocoder wurde 1965 von dem US-amerikanischen Akustikingenieur und Forscher James Flanagan und seinem Kollegen Roger Golden vorgestellt.<sup>24</sup> Etwa 10 Jahre später implementierte Michael R. Portnoff, Angestellter und Forscher der University of California, das Verfahren in digitalen Anwendungen.<sup>25</sup> Ursprünglich wurde das Verfahren zur Ver- und Entschlüsselung von Stimmen benutzt, was sich auch noch am Namen erkennen lässt (Voice Coder = Vocoder).<sup>26</sup> Durch technische Verbesserungen in den 1990er Jahren wurde das Prinzip immer häufiger für Time Stretching und Pitch Shifting im musikalischen Bereich angewendet.

---

<sup>21</sup> Vgl. Xenakis, Iannis (1960), zitiert nach Roads, Curtis (1996) in: The computer music tutorial, USA. S. 169.

<sup>22</sup> Vgl. Hörbeispiel Nr. 3.

<sup>23</sup> Vgl. Roads, Curtis (1996): The computer music tutorial, USA. S. 169.

<sup>24</sup> Vgl. Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2008): Handbuch der Tonstudioteknik, Band 1. 7. Auflage, München. S. 87.

<sup>25</sup> Vgl. Bernsee, Stephan (1999): Time Stretching and Pitch-Shifting of Audio Signals – An Overview, <http://blogs.zynaptiq.com/bernsee/time-pitcMicheal R.h-overview/> (Zugriff: 13.01.2019).

<sup>26</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2008): Handbuch der Tonstudioteknik, Band 1. 7. Auflage, München. S. 87.

Die Funktionsweise des Phase-Vocoders lässt sich etwas vereinfacht (eine genauere Betrachtung erfolgt im weiteren Verlauf dieser Arbeit) wie folgt erklären: Fensterweise, also in aufeinanderfolgenden, einzelnen Bereichen wird das Originalsignal betrachtet und dessen Parameter analysiert. Mathematisch betrachtet wird hier das Verfahren der Short Time Fourier Transformation angewendet (STFT). Aus den sich ergebenden Parametern der STFT wird nun eine Rekonstruktion des Signals erstellt, die sich dabei auch modifizieren lässt. Um einen Klang nun in seiner Tonhöhe zu verändern, beispielsweise um eine Oktave nach oben, wird dieser zunächst in doppelter Länge rekonstruiert, um dann mit ebenfalls doppelter Geschwindigkeit abgespielt zu werden.<sup>27</sup> Wesentliche Klangeigenschaften und Strukturen bleiben dabei erhalten.<sup>28</sup>

## 2.3 Anwendungen

Die ersten Musikcomputer, die während der 1950er und 60er Jahren eingesetzt wurden, waren sehr groß und teuer und dienten lediglich der Datenverarbeitung. Gegen Ende der 60er Jahre kamen kleinere und günstigere Modelle auf den Markt, die allerdings nur zur Steuerung analoger Geräte in einigen Tonstudios eingesetzt wurden, da sie für eine direkte Synthese oder Bearbeitung von Audiomaterial zu wenig Leistung lieferten. Erst in den 1970er Jahren wurden digitale Geräte konstruiert, die speziell auf die Verarbeitung von Audiomaterial ausgelegt und auch preisgünstiger waren.<sup>29</sup>

Francis F. Lees Firma Lexicon brachte 1972 als erste ein digital arbeitendes Gerät auf den Markt, mit dem eine von der Geschwindigkeit unabhängige Veränderung der Tonhöhe möglich war: die erste Version des Varispeech.<sup>30</sup> Das Gerät nutze das Verfahren der Granularsynthese und war anfangs für die Sprachtherapie konzipiert. In den folgenden Jahren spezialisierte sich Lexicon allerdings

---

<sup>27</sup> Vgl. Dolson, Mark (2014): The Phase Vocoder: A Tutorial, <http://www.panix.com/%7Ejens/pvoc-dolson.par> (Zugriff: 13.01.19).

<sup>28</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2008): Handbuch der Tonstudioteknik, Band 1. 7. Auflage, München. S. 87.

<sup>29</sup> Vgl. Manning, Peter (1993): Electronic & Computer Music, 2. Auflage, USA. S. 213.

<sup>30</sup> Vgl. Roads, Curtis (1996): The computer music tutorial, USA. S. 443.

immer mehr auf den professionellen Audiobereich und so kamen Folgemodelle des Varispeech auch im Rundfunk und der Musikproduktion zum Einsatz.

Nur wenig später, Mitte der 1970er Jahre, wurde mit dem Harmonizer H910 der Firma Eventide das erste kommerziell verfügbare Gerät angeboten, das digitale Tonhöhenveränderung in Echtzeit beherrschte.<sup>31</sup> Der Harmonizer erlaubte eine Manipulation der Tonhöhe in beide Richtungen um jeweils eine Oktave. Dies wurde erreicht, indem das Eingangssignal mit einer veränderten Sampling Rate wieder ausgegeben wurde. Technische Details hierzu werden im nächsten Kapitel genauer betrachtet. Bekannte Musiker wie David Bowie, Frank Zappa oder Eddie van Halen setzten die Harmonizer-Serie ein, die sich bis heute großer Beliebtheit erfreut.

1979 brach mit dem Computer Music Instrument und dem Synclavier die Zeit des digitalen Samplings an. Ab sofort konnte jedes beliebige Instrument bzw. aufgenommene Signal beispielsweise über eine Klaviatur in der gewünschten Tonhöhe abgespielt werden.<sup>32</sup> Heute gehören Sampler, sowohl als Hardware, besonders aber als Software, zum Grundinventar jedes Tonstudios und sind essentieller Bestandteil nahezu jeder Popmusik-Produktion.

---

<sup>31</sup> Vgl. Roads, Curtis (1996): The computer music tutorial, USA. S. 444.

<sup>32</sup> Vgl. ebd. S. 117 ff.



### 3 Funktionsweise und Anwendungen

In diesem Kapitel werden verschiedene Verfahren zur digitalen Audiosignalverarbeitung genauer betrachtet. Außerdem wird die Funktionsweise der beiden großen Tools zur Tonhöhenkorrektur, „Celemony Melodyne“ und „Antares Auto-Tune“, untersucht. Stellvertretend für DAW-eigene Möglichkeiten für Tonhöhenmanipulation werden des Weiteren Ableton Live 9 getestet und verschiedene Plug-Ins für Logic Pro X betrachtet (deren Ergebnisse waren bei der Veränderung von Stimmen allerdings sehr schlecht, weshalb sie im späteren Vergleich nicht mit aufgeführt werden (vgl. Kap. 4.1.1)).

#### 3.1 technische Grundlagen

Um analog erzeugte Klänge digital verarbeiten zu können, müssen diese zunächst binär dargestellt werden. Das Mikrofon stellt dabei den ersten Teil der Kette dar. Hier werden Schallschwingungen von einer dünnen Membran übernommen und in elektrische Schwingungen umgewandelt.<sup>33</sup> Dabei gibt es nochmals verschiedene physikalische Prinzipien des Schallempfangs und der Schallwandlung, deren genauere Erklärung allerdings für das weitere Verständnis nicht notwendig ist. Um das Signal nun in eine digitale, wert- und zeitdiskrete Form

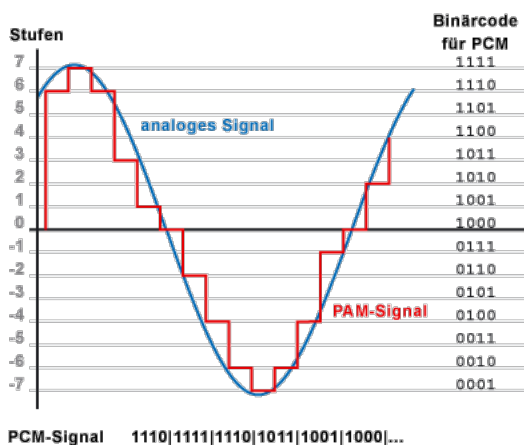


Abbildung 3: A/D-Wandlung

umzuwandeln, werden sogenannte A/D-Wandler (Analog/Digital) verwendet.<sup>34</sup> Diese Überführung passiert üblicherweise anhand der sogenannten Puls-Code-Modulation (PCM). In regelmäßigen Abständen (Abtastrate) wird die aktuelle Amplitude der Welle gemessen und auf ein digital angelegtes Raster übertragen. Aus der analogen Welle entsteht dabei das pulsamplitudenmodulierte Signal (PAM), welches sich wiederum als PCM-Signal binär darstellen lässt. Da die Zahl

<sup>33</sup> Vgl. Dickreiter, Michael (2011): Mikrofonaufnahme, 4. Auflage, Stuttgart. S. 94.

<sup>34</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2014): Handbuch der Tonstudiotechnik, Band 2, 8. Auflage, Berlin. S. 657.

der Rasterstufen systembedingt endlich ist, wird immer auf die nächstgelegene Stufe gerundet (Quantisierung). In Abbildung 3 ist dies verdeutlicht. Je nach elektrischem Bauteil können die Abtastrate (Sampling Rate) und die Abtastauflösung (Bit Rate) variieren.<sup>35</sup> Diese digitalen PCM-Signale lassen sich verlustfrei speichern bzw. formen und nachbearbeiten.<sup>36</sup>

Um das digital gespeicherte Signal nun wieder abzuspielen, spielt sich der gerade beschriebene Vorgang in umgekehrter Reihenfolge ab. Das digitale Signal wird also wieder in ein analoges umgewandelt. Bevor das geschieht, also zwischen A/D- und D/A-Wandlung, kann das Audiosignal digital manipuliert werden. Neben vielen weiteren Anwendungen kann es zerschnitten, neu zusammengesetzt, zeitlich verlängert bzw. verkürzt oder eben in seiner Tonhöhe verändert werden.

Die einfachste Form, ein Audiosignal in seiner Tonhöhe zu manipulieren, ist die gleichzeitige Veränderung der Geschwindigkeit. Das kann einerseits erreicht werden, indem man die Abtastrate des Digital-to-Analog Converters (DAC) verändert, wodurch das Signal mit einer anderen Abtastrate ausgelesen wird, als es eingelesen wurde. Ein anderes Verfahren, Sample Rate Conversion genannt, führt eine erneute Analyse des Signals (Resampling) innerhalb des Samplers durch, wobei auch hier die Abtastrate verändert wird. Allerdings passiert das noch vor dem DAC, dessen Abtastrate unangetastet bleibt.<sup>37</sup>

Um die Tonhöhe nun unabhängig von der Dauer eines Klanges zu verändern, werden in den meisten Anwendungen Verfahren verwendet, dessen Funktionsweisen bereits kurz erläutert wurden und die nun genauer untersucht werden sollen: das Prinzip der Granularsynthese (vgl. Kap. 2.2.1) und des Phase-Vocoders (vgl. Kap. 2.2.2).

Bei der Granularsynthese wird ein Klang in tausende von „Soundpartikeln“ zerschnitten und neu zusammengesetzt. Diese so genannten „Grains“ (Klangkörner) sind typischerweise zwischen 1 und 100 ms lang.<sup>38</sup> Jedes Grain ist eine sehr

---

<sup>35</sup> Die Abtastrate bestimmt die zeitlichen Abstände zwischen den Messungen des Signals, die Abtastauflösung gibt an, in wie viele Einzelstufen das analoge Signal zerlegt wird (Beispiel 16 Bit:  $2^{16}=65536$  Abstufungen).

<sup>36</sup> Vgl. Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, Martin Wöhr (2014): Handbuch der Tonstudioteknik, Band 2, 8. Auflage, Berlin. S. 658.

<sup>37</sup> Vgl. Roads, Curtis (1996): The computer music tutorial, USA. S. 124 f.

<sup>38</sup> Vgl. ebd. S. 168.

gute Repräsentation für Klänge, weil es Informationen aus dem Zeitbereich und dem Frequenzbereich kombiniert. Aufgrund der vielen verschiedenen Hüllkurven, der Dauer und anderen Eigenschaften, die ein Grain haben kann, ergibt sich eine Vielzahl an Möglichkeiten zur Verarbeitung des Signals. Die Komplexität eines durch Granularsynthese erzeugten Sounds ergibt sich aus den ihm zugeführten Daten. Wenn  $n$  die Anzahl an Parametern für jedes Grain darstellt und  $d$  die Graindichte pro Sekunde, ergeben sich  $d \times n$  Parameter, um eine Sekunde eines Klanges zu erzeugen.<sup>39</sup>

In Bezug auf eine Veränderung der Tonhöhe wird meistens auf eine Variante der Granularsynthese zurückgegriffen, die sich „Time Domain Harmonic Scaling“ (TDHS) nennt. Dafür müssen die Grains genau die Länge einer Schwingungsperiode ihrer Grundfrequenz haben.<sup>40</sup> Die Grundfrequenz ist definiert durch die Anzahl der Schwingungsperioden pro Sekunde. Durch die Veränderung der Wiedergaberate wird nun die Tonhöhe eines Grains verändert. Es ändert sich also die Anzahl der Perioden pro Sekunde und damit auch die Grundfrequenz. Um dabei die Länge des Originalsignals nicht zu verändern, werden entsprechend Grains hinzugefügt oder entnommen.

Daraus, dass das TDHS-Verfahren von einer klaren Grundfrequenz abhängig ist, ergibt sich, dass es für eine Bearbeitung monophoner (=einstimmig) Klänge sehr gut geeignet ist. Bei der polyphonen (=mehrstimmig) Verarbeitung hingegen können sich dadurch, dass keine klare Grundfrequenz erkennbar ist, Rechenfehler ergeben.<sup>41</sup> Anders die Tonhöhenmanipulation nach dem Prinzip des Phase-Vocoders – diese lässt auch eine polyphone Bearbeitung zu.

Der Phase-Vocoder (PV) arbeitet nach dem Prinzip der Resynthese. Dabei wird der mathematische Prozess der Fouriertransformation angewendet, wonach eine Wellenform in die ihr zugrunde liegenden einzelnen Sinuswellen aufgeteilt wird. Die sich daraus ergebenden Parameter können nun reproduziert und gegebenenfalls moduliert werden. Um das auf ein sich veränderndes Signal anwenden zu können, kommt die Short Time Fouriertransformation (STFT) zum Tragen. Das

---

<sup>39</sup> Vgl. Roads, Curtis (1996): The computer music tutorial, USA. S. 170 f.

<sup>40</sup> Vgl. Bernsee, Stephan (1999): Time Stretching and Pitch-Shifting of Audio Signals – An Overview, <http://blogs.zynaptiq.com/bernsee/time-pitch-overview/> (Zugriff: 12.01.2019).

<sup>41</sup> Vgl. Bernsee, Stephan (1999): Time Stretching and Pitch-Shifting of Audio Signals – An Overview, Link siehe oben (Zugriff: 12.01.2019).

Eingangssignal wird dabei in aufeinanderfolgenden Fenstern festgelegter Länge betrachtet. In der Theorie analysiert der PV also das ankommende Signal und erstellt eine Sample-by-sample-Kopie.<sup>42</sup> Durch die eingestellten Parameter der Analyse kann es allerdings in der Praxis zu Informationsverlust kommen. Diese Parameter müssen für die Analyse und die Erwartungen an die Resynthese eines Klanges entsprechend angepasst werden.<sup>43</sup> Darunter zählen beispielsweise die Anzahl der zu analysierenden Samples in einem bestimmten Zeitrahmen oder die Art des Fensters, in dem das Signal betrachtet wird. Ein Fenster ist also in diesem Fall ein bestimmter Rahmen, der den in sich befindlichen Teil des Signals analysiert. Wenn das Signal nicht gerade einen Nulldurchlauf am Rand des Fensters hat, was in den wenigsten Fällen der Fall ist, kommt es an den Übergängen zwischen zwei Fenstern zu Verzerrungen. Über kurze Fade-Outs bzw. Fade-Ins können diese reduziert werden.

Um nun ein Signal gesamtheitlich in seiner Tonhöhe um einen Faktor  $x$  zu verändern, sind zwei Schritte notwendig. Zunächst wird jede einzelne Frequenz des zu verändernden Signals mit dem Faktor  $x$  multipliziert, wodurch sich logischerweise ihre Dauer verändert. Das daraus resultierende Signal wird nun mit einer angepassten Abtastrate erneut analysiert (Resampling), wobei sich die angepasste Abtastrate aus dem ursprünglichen Wert multipliziert mit dem Faktor  $x$  ergibt. Das Ausgangssignal hat somit dieselbe Länge wie das Eingangssignal, seine Tonhöhe wurde allerdings um den Faktor  $x$  verändert.<sup>44</sup> Dadurch ist auch eine polyphone Bearbeitung des Signals möglich, da jede Frequenz einzeln moduliert wird. Ein weiterer Unterschied zum TDHS-Verfahren der Granularsynthese besteht darin, dass der PV nicht in Echtzeit, also beispielsweise bei Live-Konzerten, angewendet werden kann.

Im Folgenden werden nun verschiedene aktuelle Programme zur Tonhöhenmanipulation vorgestellt. Welches genaue technische Verfahren einem Programm dabei zugrunde liegt, ist in den meisten Fällen ein Geschäftsgeheimnis des Herstellers.

---

<sup>42</sup> Vgl. Portnoff (1976) in: Roads, Curtis (1996): The computer music tutorial, USA. S. 566.

<sup>43</sup> Vgl. Roads, Curtis (1996): The computer music tutorial, USA. S. 567.

<sup>44</sup> Vgl. Dolson, Mark; Laroche, Jean (1999): New phase-vocoder techniques for pitch-shifting, harmonizing and other exotic effects, <http://www.ee.columbia.edu/~dpwe/papers/LaroD99-pvoc.pdf> (Zugriff: 15.01.2019).

## 3.2 Antares Auto-Tune

Im Jahre 1997 kam mit dem Auto-Tune Pitch Correcting Plug-In des Herstellers Antares ein bahnbrechendes neues Tool zur Manipulation von Audiosignalen, insbesondere im Bereich der Tonhöhe, auf den Markt.<sup>45</sup> Das Programm ermöglicht es, die Stimme und andere Solo-Instrumente in Echtzeit in seiner Tonhöhe zu verändern, ohne dass dabei der Klangcharakter des Originalsignals zu stark verändert wird. Das Recording Magazine nannte Auto-Tune den „holy grail of recording“.<sup>46</sup> In den folgenden Jahren hat sich Auto-Tune als weltweiter Standard etabliert und wird täglich von vielen Tausend Nutzern im Amateur- und im professionellen Bereich genutzt. Der Produzent Ryan Shanahan, der unter anderen mit den Sängerinnen Selena Gomez und Lady Gaga zusammengearbeitet hat, wird auf der Website des Herstellers Antares mit folgendem Satz zitiert:

*„Auto-Tune is a tool I use on pretty much every vocal I engineer. I can't remember the last time I worked on a record without it.“<sup>47</sup>*

Auch der sehr bekannte, roboterartige Stimmeeffekt, wie er 1998 das erste Mal im Song „Believe“ von Cher<sup>48</sup> zu hören war und der seitdem oft als „Cher-Effekt“ bezeichnet wird, ist durch die Anwendung von Auto-Tune entstanden.

Das Tool ist für die Verwendung auf monophone Schallsignale wie die Einzelstimme oder ein einstimmiges Solo-Instrument konzipiert. Mehrere Stimmen bzw. Instrumente können nicht akkurat verändert werden.<sup>49</sup>

Antares bietet aktuell verschiedene Versionen von Auto-Tune an, die für unterschiedliche Anwendungsbereiche optimiert sind. Dazu gehören Auto-Tune Pro, Auto-Tune 7, Auto-Tune Live und Auto-Tune EFX 3.

---

<sup>45</sup> Vgl. Antares Audio Technologies (2016): Auto-Tune Live. Owner's Manual. S. 9, [https://www.antarestech.com/mediafiles/documentation\\_records/10\\_Auto-Tune\\_Live\\_Manual.pdf](https://www.antarestech.com/mediafiles/documentation_records/10_Auto-Tune_Live_Manual.pdf) (Zugriff: 16.01.2019).

<sup>46</sup> Vgl. ebd.

<sup>47</sup> Vgl. Ryan Shanahan, zitiert nach: Antares Audio Technologies (2019): <https://www.antarestech.com/product/auto-tune-pro/> (Zugriff: 16.01.2019).

<sup>48</sup> Vgl. Hörbeispiel Nr. 6.

<sup>49</sup> Vgl. Antares Audio Technologies (2018): Auto-Tune Pro. User Guide. S. 10, [https://www.antarestech.com/mediafiles/documentation\\_records/Auto-Tune\\_Pro\\_Manual.pdf](https://www.antarestech.com/mediafiles/documentation_records/Auto-Tune_Pro_Manual.pdf) (Zugriff: 16.01.2019).

### 3.2.1 Auto-Tune Pro

Auto-Tune Pro ist die neueste und umfangreichste Version des Programms. Sie kombiniert zwei unterschiedliche Modi zur Bearbeitung des Signals: den Auto Mode für Tonhöhenverschiebung und Effekte in Echtzeit und den Graph Mode für Anpassungen der Tonhöhe und Dauer nach einer Aufnahme per grafischem Editor.<sup>50</sup>

Der Auto Mode ist konzipiert für den Einsatz in Echtzeit, sei es im Studio oder auf der Bühne. Dabei wird jedes eingehende Signal automatisch auf den nächstliegenden Ton der eingestellten Tonart geschoben, wobei auch festgelegt werden kann, welche Töne der jeweiligen Tonart genutzt bzw. nicht genutzt werden sollen. Diese und weitere einzustellenden Parameter wie Retune Speed oder Humanize lassen sich per MIDI-Controller, per Automation durch die Digital Audio Workstation (DAW) oder durch einfaches Eingeben der gewünschten Werte im Vorfeld steuern.<sup>51</sup>

In Abbildung 4 ist die Basis-Ansicht des Auto Mode gezeigt. Der Retune Speed kontrolliert die Geschwindigkeit, in der ein ankommendes Signal auf einen entsprechenden Ton gepitcht wird, wobei dieser Effekt umso hörbarer bzw. unnatürlicher



Abbildung 4: Basis-Ansicht Auto Mode

wird, je weniger Zeit man Auto-Tune lässt. Die Flex-Tune- und Humanize-Parameter verändern das Signal im Sinne der Natürlichkeit und sollen das Signal „menschlicher“ klingen lassen. Zusätzlich kann anhand des Natural Vibratos ein Vibrato erzeugt bzw. stark moduliert werden. Indem man in der oberen linken

<sup>50</sup> Vgl. Antares Audio Technologies (2018): Auto-Tune Pro. User Guide. S. 8, [https://www.antarestech.com/mediafiles/documentation\\_records/Auto-Tune\\_Pro\\_Manual.pdf](https://www.antarestech.com/mediafiles/documentation_records/Auto-Tune_Pro_Manual.pdf) (Zugriff: 16.01.2019).

<sup>51</sup> Vgl. ebd. S. 11.

Ecke „Advanced“ anklickt, wird die Ansicht um weitere Einstellungsmöglichkeiten erweitert, beispielsweise den Tonarteditor.<sup>52</sup>

Durch Betätigen des Schalters oben in der Mitte lässt sich in den Graph Mode umschalten. Dieser ist für eine detaillierte Bearbeitung des Materials nach einer



Aufnahme ausgelegt. Abbildung 5 zeigt die Ansicht des Graph Mode. Nachdem das Audiosignal eingespielt wurde („Track Pitch oder Track Pitch + Time“), erscheint in dunklem Grau

Abbildung 5: Ansicht Graph Mode

dessen Wellenform. Die farbigen Kurven können entweder automatisch anhand des Buttons „Make Curves“ oder per Hand erstellt werden. Dabei repräsentiert die rote Kurve die Originaltonhöhe, die blaue Kurve die Zieltonhöhe und die grüne Kurve die aus Rot und Blau resultierende Tonhöhe.<sup>53</sup> Als andere Möglichkeit der Bearbeitung lassen sich mit dem „Note Tool“ oder durch den Button „Make Notes“ einzelne Notenblöcke erzeugen, die innerhalb des Rasters verschiebbar sind und über die gesamte Dauer der Note die gewünschte Tonhöhe halten. Diese Dauer ist anhand des „Move Point“ und „Move Region“ Tool allerdings auch editierbar. Die Verarbeitung erfolgt dabei non-destruktiv, da hier mit einer Kopie des Originalsignals gearbeitet wird.<sup>54</sup>

Weiterhin gibt es noch einige globale Parameter, die sowohl im Auto Mode als auch im Graph Mode Verwendung finden. So lassen sich beispielsweise im Feld „Input Type“ verschiedene Optionen wie Soprano, Alto/Tenor, Instrument usw. auswählen, deren jeweiliger Algorithmus für das äquivalente Audiosignal optimiert ist. Die Parameter „Key“ und „Scale“ legen die Tonlage und Tonart fest, in

<sup>52</sup> Vgl. Antares Audio Technologies (2018): Auto-Tune Pro. User Guide. S. 11 f, [https://www.antarestech.com/mediafiles/documentation\\_records/Auto-Tune\\_Pro\\_Manual.pdf](https://www.antarestech.com/mediafiles/documentation_records/Auto-Tune_Pro_Manual.pdf) (Zugriff: 16.01.2019)

<sup>53</sup> Vgl. ebd. S. 14 ff. (Zugriff: 17.01.2019).

<sup>54</sup> Vgl. ebd. S. 16 f.



der das zu veränderte Signal bearbeitet werden soll. Diese können allerdings auch anhand des Plug-Ins „Auto-Key“ automatisch erkannt werden, welches in der aktuellen Programmversion erstmalig in Auto-Tune Pro implementiert wurde. Weitere Funktionen wie „Throat“ und „Formant“ lassen eine Stimme trotz Höhenkorrektur weiterhin natürlich klingen, „Transpose“ bietet die Möglichkeit für eine weitere, globale Höhenveränderung (+/- eine Oktave) und der „Detune“ Parameter bestimmt die Referenzfrequenz (normalerweise A = 440 Hz).<sup>55</sup>

Eine Besonderheit ist der „Classic“ Button, der, wenn aktiviert, den Sound der früheren Version Auto-Tune 5 simuliert. Dieser ist besonders beliebt und wurde sehr häufig bei Pop-Produktionen verwendet. Der Nachteil hierbei ist allerdings, dass Parameter wie Formant, Throat oder Time Editing, die in späteren Versionen implementiert wurden, dadurch deaktiviert werden und nicht mehr einstellbar sind.<sup>56</sup>

### 3.2.2 Auto-Tune Live

Ein anderes Produkt von Antares, Auto-Tune Live, ist darauf spezialisiert, mit geringstmöglicher Latenz Tonhöhenkorrekturen in Echtzeit vorzunehmen. Es ist also eigens für den Einsatz bei Live-Konzerten bzw. während einer Aufnahme konzipiert und bietet selbst keine Möglichkeit zur Nachbearbeitung. Wie auch im Auto Mode der Auto-Tune Pro Version, können die Parameter per MIDI Controller angesteuert und verändert werden. So ist beispielsweise auch ein Tonartwechsel während eines Songs mit weiterhin aktiver Tonhöhenkorrektur sehr einfach realisierbar.<sup>57</sup>

Die Funktionsweise von Auto-Tune Live lässt sich wie folgt beschreiben: Innerhalb von wenigen Millisekunden erkennt das Programm die Tonhöhe des Eingangssignals. Dazu wird die Zeit zwischen den Wiederholungen der Wellenform gemessen und dadurch die Grundfrequenz bestimmt. Innerhalb der eingestellten Skala wird das Signal nun in Richtung der entsprechend nächstgelegenen Note

---

<sup>55</sup> Vgl. Antares Audio Technologies (2018): Auto-Tune Pro. User Guide. S. 18 ff, [https://www.antarestech.com/mediafiles/documentation\\_records/Auto-Tune\\_Pro\\_Manual.pdf](https://www.antarestech.com/mediafiles/documentation_records/Auto-Tune_Pro_Manual.pdf) (Zugriff: 17.01.2019)

<sup>56</sup> Vgl. ebd. S. 19.

<sup>57</sup> Vgl. Antares Audio Technologies (2016): Auto-Tune Live. Owner's Manual. S. 9, [https://www.antarestech.com/mediafiles/documentation\\_records/10\\_Auto-Tune\\_Live\\_Manual.pdf](https://www.antarestech.com/mediafiles/documentation_records/10_Auto-Tune_Live_Manual.pdf) (Zugriff: 17.01.2019).



gepitch (wie stark die Korrektur ausfällt, wird durch die Parameter „Retune Speed“ und „Humanize“ bestimmt). Klänge mit periodischem Charakter wie die Stimme oder ein einzelnes Solo-Instrument werden demnach von Auto-Tune Live schnell erkannt und können problemlos manipuliert werden, während beispielsweise eine Gruppe von Violinen, die zwar alle den gleichen Ton spielen, aber untereinander unvermeidlich leicht verstimmt sind, keine periodische Wellenform bilden und somit von Auto-Tune Live nicht verarbeitet werden können.<sup>58</sup>

Wie bereits kurz in Kapitel 2.1.3 beschrieben, sind für den Charakter eines Klanges besonders die Formanten, die abhängig von den physischen Eigenschaften einer Schallquelle sind, von großer Bedeutung. Bei der Stimme zählen hierzu die Form der Stimmbänder, die Größe des Rachens, des Mundraumes usw. Wenn nun ein Klang über große Intervalle verändert wird, ändern sich auch seine Formanten und dementsprechend die Klangidentität. Die Funktion „Formant Correction“ von Auto-Tune Live soll dem entgegenwirken.<sup>59</sup>



Abbildung 6: Ansicht Auto-Tune Live

In Abbildung 6 ist die Oberfläche von Auto-Tune Live gezeigt. Die Parameter in der oberen Zeile entsprechen in etwa denen von Auto-Tune Pro. Das Fenster in der Mitte zeigt alle Halbtöne einer Oktave an und bietet die Möglichkeit, entweder anhand der Buttons rechts

schnell zwischen Moll und Dur umzuschalten, oder links die Halbtöne individuell anzupassen. Dafür stehen die Spalten Bypass und Remove zur Verfügung. Die ausgewählten Töne in der Remove-Spalte werden aus der aktuellen Skala entfernt und werden von Auto-Tune als Korrekturnote nicht mehr angesteuert, die

<sup>58</sup> Vgl. Antares Audio Technologies (2016): Auto-Tune Live. Owner's Manual. S. 10, S. 12 f, [https://www.antares-tech.com/mediafiles/documentation\\_records/10\\_Auto-Tune\\_Live\\_Manual.pdf](https://www.antares-tech.com/mediafiles/documentation_records/10_Auto-Tune_Live_Manual.pdf) (Zugriff 18.01.2019).

<sup>59</sup> Vgl. ebd. S. 15.

Korrektur springt also auf den nächstliegenden, in der Skala befindlichen Ton. Wenn bei einem Notenwert der Bypass-Button aktiviert ist, wird jedes Eingangssignal, das sich in der Nähe dieses Wertes befindet, ohne Korrektur durchgeschleift und unverändert wiedergegeben. Dies findet beispielsweise Anwendung, wenn nur ein bestimmter Bereich während einer Performance moduliert werden soll.<sup>60</sup> Der schwarze Balken oberhalb dieses Fensters zeigt an, um wie viele Cents und in welche Richtung das Eingangssignal gepitcht wird, wobei 100 Cents einem Halbton entsprechen.

Links neben dem Fenster befinden sich die bereits weiter oben beschriebenen Parameter (vgl. Kap. 3.2.1), während auf der rechten Seite verschiedene Einstellungen für ein künstlich generiertes Vibrato zu finden sind. Ganz unten befindet sich das „Virtual Keyboard“. Dieses erkennt die Tonhöhe des eingehenden Signals und stellt diese auf der Klaviatur dar. Des Weiteren können hier die aktuellen Tonart-Einstellungen angezeigt und das Verhalten bestimmter Noten in einer spezifischen Oktave verändert werden.<sup>61</sup>

### 3.2.3 Auto-Tune EFX 3 und Auto-Tune 7

Auto-Tune EFX 3 ist eine Version von Auto-Tune, die den „Auto-Tune Vocal Effect“ (weiter oben als „Cher-Effekt“ bezeichnet) bei einfacher Bedienbarkeit verfügbar machen soll. Auch diese deutlich preisgünstigere Version arbeitet in Echtzeit und bietet in begrenzter Form die Möglichkeiten von Auto-Tune Live. Der Auto-Tune Vocal Effect lässt sich technisch als „Tonhöhen-Quantisierung“ bezeichnen. Dabei fallen die kleinen Variationen und fließenden Übergänge in der Stimme weg, da jede Note auf einen exakten Notenwert festgelegt ist und keine noch so kleine Abweichung davon erlaubt.<sup>62</sup>

Ein weiterer Effekt, der vom Hersteller als „Auto-Motion Vocal Pattern Generation“ bezeichnet wird, generiert musikalische Muster anhand einer Stimme. Es

---

<sup>60</sup> Vgl. Antares Audio Technologies (2016): Auto-Tune Live. Owner’s Manual. S. 28, [https://www.ataretech.com/mediafiles/documentation\\_records/10\\_Auto-Tune\\_Live\\_Manual.pdf](https://www.ataretech.com/mediafiles/documentation_records/10_Auto-Tune_Live_Manual.pdf) (Zugriff: 18.01.2019).

<sup>61</sup> Vgl. ebd. S. 30 f.

<sup>62</sup> Vgl. Antares Audio Technologies (2016): Auto-Tune EFX 3, Owner’s Manual, S. 4, [https://www.ataretech.com/mediafiles/documentation\\_records/11\\_Auto-Tune\\_EFX\\_3\\_Manual.pdf](https://www.ataretech.com/mediafiles/documentation_records/11_Auto-Tune_EFX_3_Manual.pdf) (Zugriff: 18.01.2019).

werden also nach verschiedenen, auswählbaren Presets Melodien oder andere Effekte erzeugt und der Stimme hinzugefügt.

Die Version Auto-Tune 7 ist die Vorgängerin von Auto-Tune Pro und dieser in ihrer Funktion sehr ähnlich. Der auffälligste Unterschied ist das überarbeitete und modernere Design von Auto-Tune Pro, welches aus Sicht des Verfassers eine bessere Übersicht bietet. Auf technischer Seite fehlen der Auto-Tune 7 Version nur der „Classic Mode“ Button und das implementierte „Auto-Key“ Plug-In, die beide weiter oben erwähnt sind.<sup>63</sup>

Neben den verschiedenen Auto-Tune-Versionen bietet Antares auch noch andere Programme wie beispielsweise „Harmony Engine Evo“ oder „Choir Evo“ an, die aus monophonem Material Harmonien bzw. Mehrstimmigkeit erzeugen.<sup>64</sup>

### 3.3 Celemony Melodyne

Zur Bearbeitung von Tonhöhenstrukturen gibt es neben Auto-Tune ein zweites großes Tool: Melodyne der Firma Celemony. Entwickelt wurde die Software vom deutschen Musiker, Mathematikliebhaber, Gitarrenbauer und Harmonik-Spezialist Peter Neubäcker, der 1997 mit der Entwicklung begann. 2001 kam Melodyne erstmals als Stand-Alone<sup>65</sup> auf den Markt und ist mittlerweile bei seiner 4. Version angekommen.<sup>66</sup> Während Auto-Tune eher auf die Tonhöhenkorrektur in Echtzeit spezialisiert ist, kommt Melodyne ausschließlich bei der Nachbearbeitung einer Aufnahme zum Einsatz. Ein weiterer großer Unterschied zu Auto-Tune besteht darin, dass Melodyne anhand einer eigens entwickelten und patentierten Technologie (DNA = Direct Note Access) in der Lage ist, polyphone Klänge in ihre Ursprungs Klänge aufzuteilen und somit beispielsweise die Bearbeitung von Akkorden oder mehrstimmigem Gesang ermöglicht.<sup>67</sup> Je nachdem, welche Funktionen man für

---

<sup>63</sup> Vgl. Thornton, Mike (2018): Antares Announces Auto-Tune Pro to Replace Auto-Tune 8, <https://www.pro-tools-expert.com/home-page/2018/1/25/antares-announces-auto-tune-pro> (Zugriff: 18.01.2019).

<sup>64</sup> Vgl. Antares Audio Technologie (2019): Products, <https://www.antarestech.com/product/harmony-engine-evo/> (Zugriff: 18.01.2019).

<sup>65</sup> Melodyne besaß eine eigene Benutzeroberfläche und noch nicht in DAWs integrierbar. Seit 2007 ist Melodyne auch als Plug-In verfügbar.

<sup>66</sup> Vgl. Celemony Software GmbH (2019): Die Firma, <https://www.celemony.com/de/service1/about-celemony/company> (Zugriff: 19.01.2019).

<sup>67</sup> Vgl. ebd.

die Bearbeitung benötigt, bietet Celemony vier verschiedene Versionen an: essential, assistant, editor und studio. Im Folgenden wird nur Melodyne studio genauer betrachtet, da diese die umfangreichste aller Versionen ist und somit jede Funktion, die Melodyne bietet, enthalten ist.

In Abbildung 7 ist die die Arbeitsoberfläche von Melodyne studio zu sehen. In der oberen linken Ecke wird angezeigt, welche Bearbeitungsfenster gerade aktiv sind. In diesem Fall sind das Informationsfenster, die Spurenansicht und der Noteneditor aktiviert. Nicht zu sehen ist der Sound Editor. Das Informationsfenster, welches sich übrigens auch auf der rechten Seite öffnen lässt, enthält Informati-

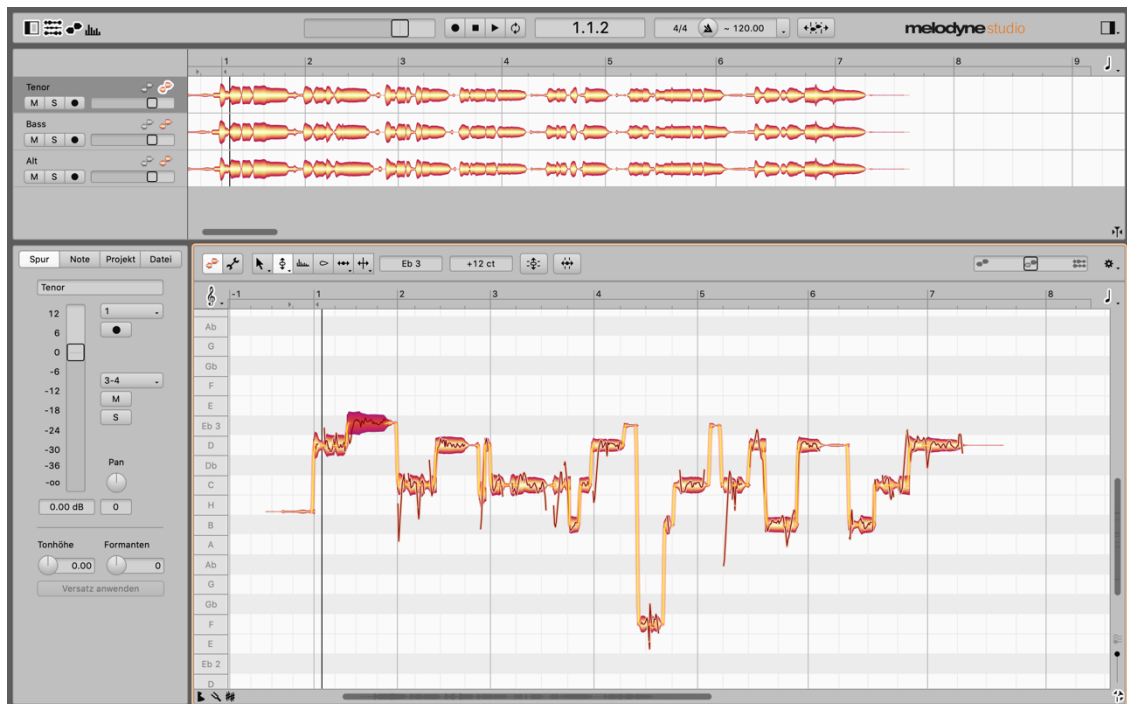


Abbildung 7: Ansicht Melodyne

onen über den jeweils ausgewählten Reiter (in diesem Fall „Spur“). Die Spurenansicht zeigt alle im Projekt befindlichen Spuren, ähnlich wie bei den meisten DAWs. Hier lässt sich auch auswählen, welche Spur man gerade bearbeiten möchte. Das dritte Symbol schaltet den Noteneditor ein bzw. aus. Hier erfolgen die wesentlichen Bearbeitungsschritte, deshalb wird dieser später genauer betrachtet. Der in diesem Fall deaktivierte Sound Editor erlaubt noch mal eine gesonderte Bearbeitung der ausgewählten Note. Man kann hier anhand eines Equalizers sehr stark in deren Frequenzspektrum eingreifen und durch Veränderungen der Obertonstruktur eine komplett andere Klangfarbe erzeugen.

Außerdem lässt sich hier ein Klang mit frei wählbar veränderten Parametern neu erzeugen (resynthesize).<sup>68</sup>

Weiter in der oberen Zeile folgt nun ein Regler für die Hauptlautstärke und Buttons für die Funktionen Record, Stopp, Start und Loop. Das kleine Fenster rechts daneben zeigt an, in welchem Takt innerhalb des Projektes man sich gerade befindet. Noch einmal weiter sind die Taktart und das Tempo einzustellen, mit einem integrierten Metronom-Button. Das Tempo lässt sich entweder manuell eingeben oder kann von Melodyne automatisch erkannt werden, sobald eine Audiospur vorhanden ist. Mit dem Auto-Stretch-Button rechts vom Tempo-Fenster werden, wenn aktiviert, alle Noten einer Audiospur automatisch dem vorgegebenen Tempo und der Taktart angepasst.

Audiomaterial kann auf verschiedene Weisen in Melodyne eingepflegt werden. Zum einen gibt es die Möglichkeit der direkten Aufnahme ins Programm, zum anderen können Audiosignale auch per Drag und Drop oder durch „Ablage/Audio importieren“ importiert werden. Sobald eine Audiospur in Melodyne angekommen ist, wird kurz gerechnet und anschließend wird das Signal, zerlegt in Notenhöhe, -länge, -amplitude, Vibrato und Übergänge zwischen den Noten, im Noteneditor dargestellt. Für die Analyse kann dabei im Vorfeld aus verschiedenen Algorithmen gewählt werden, die unterschiedlich gut für verschiedene Anwendungen wie beispielweise mono- oder polyphones Material geeignet sind.

Nun gibt es verschiedene Tools zur Bearbeitung der unterschiedlichen Parameter jeder Note.<sup>69</sup> Das Tonhöhenwerkzeug wird für das Verschieben der Noten verwendet. Dabei können eine oder mehrere Noten ausgewählt und an die gewünschte Stelle gezogen werden. Sobald eine Note ausgewählt ist, wird in einem Fenster rechts neben den Tools angezeigt, welchem Notenwert diese am nächsten ist und um wie viele Cents sie davon abweicht. Mit einer anderen Form desselben Werkzeuges lassen sich auch das Vibrato innerhalb einer Note glätten bzw. verstärken und der Übergang zwischen zwei Noten anpassen. Mit dem Formantwerkzeug können die Formanten der verschiedenen Noten verändert werden. Dadurch ist es beispielsweise möglich, einen „Chipmunk-Effekt“ zu

---

<sup>68</sup> Vgl. Celemony Software GmbH (2019): Melodyne 4, Handbuch, S. 146 ff, <http://helpcenter.celemony.com/hc-2/pdf/melodyneStudio4/de?env=standAlone> (Zugriff: 20.01.2019).

<sup>69</sup> Vgl. ebd. S. 98 ff.

erzeugen, ohne dass sich die Tonart ändert. Das Amplitudenwerkzeug findet Verwendung, sobald die Lautstärken einzelner Noten untereinander angepasst werden müssen. Dies ist auch visuell nachzuvollziehen, da mit einer Veränderung der Amplitude auch die Noten dicker oder dünner dargestellt werden. Das vorletzte Tool ist das Timingwerkzeug, womit Veränderungen in der Dauer oder Position einer Note hervorgerufen werden können. Es bietet außerdem die Möglichkeit, die Attack, also quasi den Einsatz einer Note zu manipulieren. Melodyne trennt die einzelnen Noten grundsätzlich nach Tonhöhe und starken Amplitudenschwankungen. Dementsprechend kann es vorkommen, dass zwei Noten, die falsch eingesungen wurden, als eine erkannt werden. Mit dem letzten Tool, dem Notentrennwerkzeug, lässt sich dies einfach beheben, indem man die Noten an der gewünschten Stelle voneinander trennt und somit voneinander unabhängig anpassen kann.

Des Weiteren befinden sich rechts neben der Werkzeugauswahl noch zwei Buttons, die einen schnellen Zugriff auf die Tonhöhenkorrektur und die Quantisierung aller ausgewählten Noten bieten. Bei der Tonhöhe kann bestimmt werden, zu wieviel Prozent die falschen Noten in Richtung der exakten Note gepitcht werden. Außerdem ist hier auch ein Feld anwählbar, wodurch alle Noten in der entsprechenden Tonart einrasten, die Melodyne übrigens ebenfalls automatisch erkennt. Die Intensität der Quantisierung, also wie sehr sich die ausgewählten Noten an das angegebene Taktraster anpassen, kann auf gleiche Weise anhand eines Reglers bestimmt werden.<sup>70</sup>

Wenn man letztendlich mit der Bearbeitung zufrieden ist, kann man nun, wenn man Melodyne als Stand-Alone verwendet, entweder alle Spuren einzeln oder als Stereo-Mix beispielsweise im wav-Dateiformat exportieren und in der DAW daran weiterarbeiten. Wenn Melodyne als Plug-In verwendet wird, ist es ratsam, dieses an die erste Stelle der Effektkette zu setzen, da man dadurch einerseits sehr gut vor und nach der Bearbeitung vergleichen kann und andererseits jede weitere Bearbeitung wie Equalizer, Kompressoren etc. nach der Tonhöhenkorrektur passiert.

---

<sup>70</sup> Vgl. Celemony Software GmbH (2019): Melodyne 4, Handbuch. S. 92 ff, <http://helpcenter.celemony.com/helpcenter/pdf/melodyneStudio4/de?env=standAlone> (Zugriff: 20.01.2019).



Als weiteres Feature ist noch zu nennen, dass man eine Audiospur auch als MIDI-Signal exportieren kann. Dadurch ist es z.B. möglich, eine eingesungene Melodie von jedem beliebigen anderen Instrument abspielen zu lassen.

### 3.4 Logic Pro X und Ableton Live 9

Abgesehen von den beiden Studiostandards haben die meisten DAWs eigene Plug-Ins zur Tonhöhenkorrektur. Zunächst werden Pitch Correction und Pitch Shifter von **Logic Pro X** betrachtet. In Abbildung 8 sind diese beiden Programme zu sehen. Pitch Correction funktioniert dabei ähnlich wie Auto-Tune. Es analysiert das Eingangssignal und korrigiert die falschen Töne auf den nächstgelegenen

Ton der angegebenen Skala, die sich manuell festlegen lässt. Zusätzlich können auch hier einzelne Noten mit dem Bypass-Button deaktiviert werden. Je nachdem,



Abbildung 8: Pitch Shifter und Pitch Correction (Logic Pro X)

ob „show input“ oder „show output“ angewählt ist, leuchten die entsprechenden Noten in Echtzeit auf. Die Leiste darunter zeigt dabei an, um wie viele Cents eine Note korrigiert wurde. Eine Korrektur von polyphonem Material ist übrigens nicht möglich.<sup>71</sup> Die beiden Parameter rechts neben der Klaviatur bestimmen zum einen, wie schnell eine Anpassung erfolgt, und zum anderen kann damit das gesamte Signal um bis zu 50 Cents nach oben oder unter verschoben werden. Wenn die Lampe unter „Use Global Tuning“ links neben der Klaviatur aktiviert ist, sind der Korrektur kein Grenzen gesetzt und ein falscher Ton wird gepitcht, egal,

<sup>71</sup> Vgl. Apple Inc. (2018): Logic Pro X, Pitch Correction – Übersicht, <https://help.apple.com/Logicpro/mac/10.4/#/lgcef2835dcc> (Zugriff: 21.01.2019).

wie viele Cents sie von der nächstgelegenen Note entfernt ist. Ist diese Lampe deaktiviert, lässt sich darunter bestimmen, um wie viele Cents eine Note maximal korrigiert werden soll. Eine Art Humanize-Button oder Ähnliches gibt es nicht.

Der Pitch Shifter von Logic Pro X funktioniert wie jeder andere auch. Das gesamte Signal kann um bis zu 12 Halbtöne in jede Richtung gepitcht werden. Eine Formantenkorrektur gibt es dabei nicht. Auch hier kann das Signal nochmal um jeweils 50 Cents (also einen Viertelton) nach oben oder unten verschoben werden und der Mix-Regler bestimmt, wie sehr der Effekt im Signal zu hören ist.

**Ableton Live** bietet ebenfalls eine Möglichkeit, ein Signal in seiner Tonhöhe zu verändern. Dies funktioniert hier objektbasiert. Durch doppeltes Anklicken eines Audioobjekts öffnen sich verschiedene Bearbeitungsfelder, von denen eines „Sample“ heißt. Hier kann ein Sample um bis zu 48 Halbtöne in jede Richtung transponiert werden. Sogar einen einstellbaren Parameter für die Formanten gibt es. Ein Plug-In zur automatischen Tonhöhenkorrektur gibt es bei Ableton Live 9 Suite allerdings nicht.

Diese Beispiele sollen stellvertretend für die DAW-eigenen Plug-Ins zur Tonhöhenveränderung stehen. Keines dieser Plug-Ins erreicht dabei allerdings den Umfang von Auto-Tune und Melodyne. Im nächsten Kapitel dieser Arbeit werden die Klangergebnisse und die jeweiligen Funktionsumfänge der verschiedenen Programme miteinander verglichen.



## 4 Vergleich

Jedes der oben beschriebenen Programme arbeitet auf eine andere Art und Weise. Auch die Benutzeroberfläche und dementsprechend die Übersicht und Handhabung der einzelnen Programme unterscheiden sich. Im Folgenden werden Klangergebnisse, Funktionsumfang und Handhabung der einzelnen Programme verglichen. Dabei liegt der Fokus auf dem Vergleich zwischen Antares Auto-Tune und Celemony Melodyne, da diese speziell für den Zweck der Tonhöhenmanipulation ausgelegt sind und als die beiden besten Vertreter ihrer Art gelten. Verschiedene Versionen und Varianten beider Programme wurden bereits im letzten Kapitel beschrieben. Für den Vergleich werden mit Auto-Tune Pro und Melodyne studio die jeweils umfangreichsten Versionen verwendet. Wenn im Folgenden also von Melodyne bzw. Auto-Tune gesprochen wird, sind jeweils diese Versionen gemeint, es sei denn, es wird ausdrücklich eine andere Version genannt.

Für den Vergleich arbeitete der Verfasser mit Testversionen von Melodyne und Auto-Tune. Diese liefern den vollen Funktionsumfang inklusive Speichern, Exportieren usw. mit dem einzigen Unterschied, dass die Lizenzen bei Auto-Tune nach 10 und bei Melodyne nach 30 Tagen ablaufen. Die Aktivierung erfolgte ebenso auf unterschiedliche Weise. Melodyne ließ sich sehr einfach und komfortabel online aktivieren, während selbst für die Auto-Tune Testversion ein iLok-Konto und ein iLok-USB Key<sup>72</sup> angeschafft werden mussten, welcher sich als unnötig teuer herausstellte. War dies allerdings erledigt, ging die Installation und Aktivierung reibungslos vonstatten.

Um für die Klangbeurteilung nicht nur die Meinung des Verfassers als Maßstab anzuführen, wurde eine Befragung durchgeführt. Hierfür wurden zehn Personen verschiedenen Alters und mit unterschiedlichem Bezug zur Musik befragt. Hardwaretechnisch wurden dabei immer dieselben Komponenten benutzt, um für jeden Befragten das gleiche Klangbild zu gewährleisten. Diese Umfrage ist

---

<sup>72</sup> iLok ist ein Lizenzen-Management-System, das mit verschiedenen Anbietern von Musiksoftware zusammenarbeitet und verhindern soll, dass Lizenzen illegal beschafft werden können, indem für die Aktivierung ein spezieller USB-Stick verwendet wird.

tabellarisch ausgewertet dieser Arbeit angehängt. Wichtige Ergebnisse werden an entsprechenden Stellen in den Text eingebunden.

## **4.1 Vergleich gemeinsamer Features**

Melodyne und Auto-Tune bieten unterschiedlich umfangreiche Möglichkeiten zur Bearbeitung. In diesem Teil werden zunächst Funktionen getestet und verglichen, die beide Programme gemeinsam haben. Spezielle Funktionen, wie die polyphone Bearbeitung bei Melodyne oder die Möglichkeit zum Live-Einsatz bei Auto-Tune werden später gesondert angeführt (vgl. Kap. 4.2).

Verglichen wurden drei verschiedene Aspekte, die nachfolgend in einzelne Teile gegliedert sind: Im ersten Teil wurde ein mehrstimmiges Lied von vier Sängern und Sängerinnen eingesungen, wobei jede Stimme mit beiden Programmen einzeln in ihrer Tonhöhe korrigiert und dann als Ganzes zusammengesetzt wurde. Die Befragten sollten danach beurteilen, welches Ergebnis besser intoniert war und harmonischer klang. Im zweiten Teil wurde eine einstimmig gesungene Melodie mit verschiedenen Programmen halbtöne verschoben, deren Ergebnisse anschließend beurteilt wurden. Im dritten Teil wurde versucht, den Klingeffekt aus Chers Hit „Believe“ (bekannt als „Cher- oder Auto-Tune-Effekt“) zu simulieren.

### **4.1.1 Tonhöhenkorrektur**

Der erste Test erfolgte mit einer Aufnahme des Kirchenlieds „Amazing Grace“ in einem vierstimmigen Arrangement. Zu den Sängern und Sängerinnen zählten ein Chorleiter, eine Berufssängerin, ein Berufssänger und eine langjährige Hobby-Chorsängerin. Jede Stimme wurde einzeln eingesungen und einer eigenen Spur zugeordnet. Aufgenommen wurde mit einem Großmembran-Kondensatormikrofon, welches per Focusrite-Interface angeschlossen war. Als Aufnahme- und Host-DAW wurde Logic Pro X verwendet.

Untersuchungsziel war, zu vergleichen, inwieweit sich die Ergebnisse von Auto-Tune und Melodyne unterscheiden. Dafür wurde jede Stimme jeweils mit beiden Programmen korrigiert und als Arrangement zusammengeführt. Hörbeispiel Nr.

12 ist die Version ohne Korrektur, wodurch die Unterschiede zwischen vorher und nachher verdeutlicht werden sollen.<sup>73</sup>

Ursprünglich sollte auch das Logic eigene Pitch Correction Tool getestet werden. Die Ergebnisse fielen dabei aber so schlecht aus, dass ein Vergleich mit Melodyne und Auto-Tune wenig Sinn ergeben hätte.

### **4.1.1.1 Melodyne**

Für diesen Test wurde die Plug-In-Variante von Melodyne verwendet. Die Stand-Alone-Variante wird weiter unten nochmal erwähnt. Da mehrere Instanzen des Programms auf unterschiedliche Spuren geladen werden, wird der in Kapitel 3.3 beschriebene Spureneditor im Ansichtsfenster durch eine Spurenliste ersetzt, die alle Spuren auflistet, die mit einer Melodyneinstanz belegt sind. Das Öffnen vieler verschiedener Instanzen, insbesondere in Verbindung mit weiteren Effekten auf den einzelnen Spuren, kann dabei je nach Leistung des Computers zu Performanceeinschränkungen führen, da Melodyne ziemlich viel Rechenleistung beansprucht. Diese Problematik ist allerdings mit einfachen Tricks wie „Einfrieren“ der Spuren zu umgehen.

Mit der Erkennung der Notenwerte hatte Melodyne keinerlei Probleme. Dabei ist beim Transferieren einer Spur der zuvor eingestellte Algorithmus für die Erkennung sehr wichtig. Insbesondere die beiden Berufssänger hatten ein starkes Vibrato in ihren Stimmen, welches besonders bei lang gehaltenen Tönen stark ausgeprägt war. Beim ersten Transferieren war versehentlich der polyphone Algorithmus eingestellt (vgl. Kap. 4.2.1.2), sodass eben diese Vibratos als unterschiedliche Tonhöhen angezeigt wurden, die eine natürlich klingende Korrektur sehr schwierig gestalteten. Durch einfaches Wechseln des Algorithmus und Neuberechnung des Materials verwandelten sich aus mehreren Noten eine einzelne mit einer sehr starken Modulationskurve (vgl. Abb. 9).

---

<sup>73</sup> Vgl. Hörbeispiel Nr. 12.

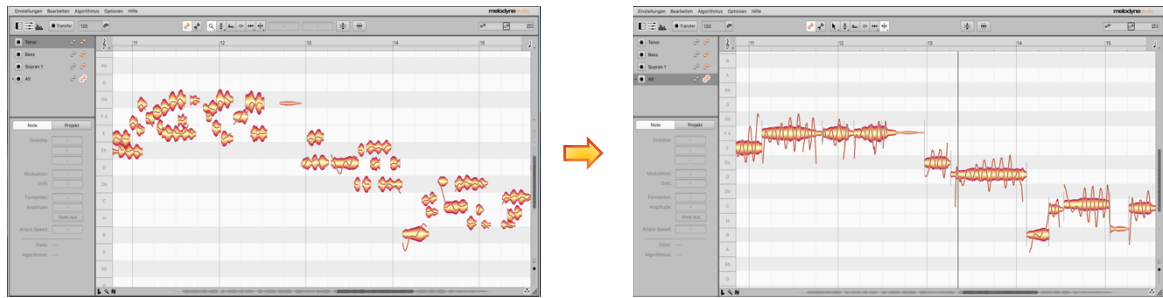


Abbildung 9: polyphoner und monophoner Algorithmus

Nachdem also der richtige Algorithmus ausgewählt war, wurde jede Spur nach dem Transferieren richtig erkannt und in der jeweiligen Instanz dargestellt. Es erwies sich als sinnvoll, immer nur eine Instanz zur gleichen Zeit geöffnet zu haben, da es ansonsten wie bereits beschrieben zu Performanceeinschränkungen oder Programmabstürzen kommen konnte. Es ist sogar möglich, innerhalb einer Instanz zwischen den zu bearbeitenden Spuren zu wechseln, allerdings kam es auch dabei hin und wieder zu Verwirrungen innerhalb des Programms. Die sicherste Methode ist also, für jede Spur die entsprechende Instanz einzeln zu öffnen und nach der Bearbeitung wieder zu schließen. Wenn man nun die restlichen Spuren zusätzlich zu der gerade geöffneten hören will, gibt es in der oberen rechten Ecke des Noteneditors einen Balken, der, je weiter rechts sich der Cursor befindet, immer lauter alle anderen mit Melodyne belegten Spuren dem Signal zuführt. Wenn sich der Cursor also ganz rechts befindet, werden alle Spuren gleichlaut gehört. Außerdem ist es möglich, sich die Notenwerte aller Spuren gleichzeitig im Noteneditor anzeigen zu lassen, an denen man sich gut orientieren kann.

Nun gibt es verschiedene Möglichkeiten zur Korrektur. Zum einen lässt sich jede Note einzeln auf die gewünschte Position verschieben oder kann bei ausgewähltem Tonhöhenwerkzeug per Doppelklick exakt auf den nächstgelegenen Halbtton korrigiert werden. Zum anderen können alle Noten gleichzeitig markiert und entweder näherungsweise oder zu hundert Prozent in Richtung des nächsten Halbtons verschoben werden.

### 4.1.1.2 Auto-Tune

Ein zweites Hörbeispiel wurde mit Auto-Tune erstellt. Um die Vergleichbarkeit zu gewährleisten, wurden selbstverständlich dieselben Aufnahmen verwendet wie

bei der Bearbeitung mit Melodyne. Aus diesem Grund arbeitete Auto-Tune zwar in Echtzeit, allerdings mit dem bereits aufgenommenen Material und nicht live während der Aufnahme.

Für die Bearbeitung wurde nach demselben Verfahren wie mit Melodyne jede Stimme mit einer eigenen Instanz von Auto-Tune belegt und einzeln korrigiert. Auch hier wurde immer nur eine Instanz zur selben Zeit geöffnet, Performanceeinschränkungen gab es dabei keine.

Jedes Eingangssignal wird von Auto-Tune bearbeitet, sobald eine Instanz im Kanalzug der Spur aktiviert wurde. Das Programm fängt also direkt an zu arbeiten und unterscheidet sich in diesem Punkt von Melodyne, welches das Material ohne die direkte Bearbeitung durch den Anwender zunächst unangetastet lässt.

Wie bereits in Kapitel 3.2.1 beschrieben, richtet sich Auto-Tune bei der Korrektur nach der eingestellten Tonart. Im Default (Grundeinstellungen) wird dafür eine chromatische Tonleiter verwendet, es wird also jeder Halbton als mögliches Ziel der Korrektur festgelegt. Zur Bestimmung dieser wurde auf das in der Auto-Tune Pro Lizenz enthaltene Plug-In Auto Key zurückgegriffen, das die Tonart fehlerfrei erkannte und jede Instanz mit dieser Information versorgte (vgl. Kap. 3.2.1). Die Einstellungen der anderen Parameter für die Korrektur wurden nach den Einschätzungen des Verfassers so angepasst, dass eine Korrektur stattfindet, die etwaige falsche Töne auf der einen Seite korrigiert, den Klangcharakter auf der anderen Seite allerdings erhält. Als eine der wichtigsten Einstellungen stellte sich dabei der Retune-Speed heraus, der letztendlich für diesen Test mit einem Wert von 23 eingestellt wurde. Nachdem der Verfasser mit dem Klang auf einer Spur zufrieden war, wurden die Einstellungen auf die Instanz der nächsten Spur übertragen und nur noch Kleinigkeiten (z.B. die Änderung des Eingangssignals von Alt zu Tenor) angepasst.

Zeitlich gelang der Verfasser mit Auto-Tune etwas schneller zum Ziel. Hier wurden der Großteil der Parameter übertragen, sobald diese einmal eingestellt waren, während bei Melodyne die Korrektur für jede der vier Stimmen einzeln vorgenommen werden musste.

#### **4.1.1.3 Ergebnis**

Nun wurden die einzelnen Stimmen als ein Arrangement zusammengeführt und noch etwas im Panorama verteilt. Außerdem wurde jede Stimme abgesehen von der Tonhöhenkorrektur noch mit einem Kompressor, Equalizer und etwas Hall belegt. Die exakt gleichen Effekte liegen selbstverständlich auch auf der Version ohne Tonhöhenkorrektur, um die Vergleichbarkeit zu gewährleisten.

Die Ergebnisse beider Programme weisen aus Sicht des Verfassers kaum hörbare Unterschiede auf. Diese Einschätzung spiegelt sich auch in der Umfrage wider. Die Fragestellung an die Teilnehmenden war, welches der Beispiele sie als besser intoniert und insgesamt harmonischer empfanden. Fünf der zehn Befragten entschieden sich für das mit Melodyne korrigierte Beispiel, vier Personen entschieden sich für die Auto-Tune Variante und einer konnte keine Unterschiede feststellen. Alle Hörbeispiele und die Ergebnisse der Umfrage finden sich im Anhang.<sup>74</sup>

#### **4.1.2 Tonhöhenverschiebung**

Für den nächsten Test wurden die ersten beiden Zeilen des bekannten Kultliedes „I Am Sailing“ von Rod Stewart aufgenommen und anschließend als Ganzes in seiner Tonhöhe um einen Halbton nach unten verschoben. Es sollte getestet werden, inwiefern sich die mit verschiedenen Programmen manipulierten Aufnahmen klanglich unterscheiden. Eine Tonhöhenkorrektur der einzelnen Noten an sich wurde dabei nicht vorgenommen, es ging allein um die Veränderung der Tonart.

Für die Aufnahme wurden dieselben technischen Komponenten genutzt wie auch schon bei der Amazing-Grace-Aufnahme.

Um sich bei der Aufnahme besser orientieren zu können, wurde eine einfache Karaoke-Version des Songs zurechtgeschnitten und als Backingtrack verwendet. Für die entsprechende Transponierung dessen wurde der Algorithmus von Ableton Live 9 Suite verwendet, da dieser für fertiges Soundmaterial ziemlich gut funktioniert. Aufgrund dieser Tatsache entschied sich der Verfasser, für den

---

<sup>74</sup> Vgl. Hörbeispiele Nr. 13-14.

Vergleich neben Auto-Tune und Melodyne auch dieses Programm für die Verschiebung der Stimme zu testen und in den Vergleich mit einzubeziehen.

Da für den Erhalt des Klangcharakters einer Stimme die Formanten besonders wichtig sind (vgl. Kap. 2.1.3), wird darauf innerhalb jedes Programms nochmal kurz eingegangen.

### **4.1.2.1 Melodyne**

Die Verschiebung aller Notenwerte gleichzeitig ist mit wenigen Handgriffen erledigt. Nach dem Transferieren des Audiosignals im richtigen Algorithmus wird jede Note wie bereits beschrieben im Noteneditor dargestellt (vgl. Kap. 3.3). Anhand des Befehls „Alle markieren“ (Strg+A für PC, cmd+A für Mac) sind alle Noten ausgewählt und lassen sich nun mit dem Tonhöhenwerkzeug einfach auf den gewünschten Ton ziehen. Dabei rastet die Bearbeitung bei jedem Halbtonschritt ein. Mit gedrückter Alt-Taste ist dieses Raster aufgehoben und die Noten lassen sich an jede beliebige Position auch zwischen den Halbtönen ziehen.

Bei der Verschiebung der einzelnen Noten selbst versucht Melodyne, die Formanten so wenig wie möglich zu verändern, damit der Charakter des Klanges erhalten bleibt. Sollte man die Formanten doch verändern wollen, was in kleinem Maße durchaus sinnvoll sein kann, um die Stimme höhen- oder tiefenlastiger klingen zu lassen, lässt sich dies entweder anhand eines Rädchens in der Info-Zeile der Spur erreichen oder anhand des Formantwerkzeuges, dessen starker Einsatz allerdings sehr schnell sehr unnatürlich klingt. Die Bedienung dieser und anderer Werkzeuge ist dabei sehr intuitiv und man gelangt ziemlich schnell in einen Workflow. Darauf wird weiter unten im Kapitel 4.3 noch mal genauer eingegangen.

### **4.1.2.2 Auto-Tune**

Mit Auto-Tune ist eine Verschiebung der Stimmlage auf verschiedene Weisen möglich. Die einfachste Möglichkeit bietet ein „Transpose“ Rädchen in der oberen rechten Ecke des Bearbeitungsfensters. Dies ist mit wenigen Klicks erledigt und nimmt dementsprechend auch sehr wenig Zeit in Anspruch. Eine zweite

Möglichkeit ist das Ziehen von Notenlinien im Graph Mode (vgl. Kap. 3.2.1). Diese Variante ist allerdings deutlich unkomfortabler und auch die Erhaltung der Natürlichkeit in der Stimme gestaltet sich dabei schwieriger. Ein Auswählen und einfaches Verschieben aller Notenwerte, so wie es mit Melodyne möglich ist, funktioniert mit Auto-Tune nicht, ist allerdings für die Verschiebung aller Noten um den gleichen Wert aufgrund der einfachen Transpose-Funktion auch nicht notwendig.

Auch bei Auto-Tune gibt es für die Bearbeitung der Formanten eigene Einstellungen. Ein Button in der oberen rechten Ecke der Benutzeroberfläche kann eine Bearbeitung de- und aktivieren. Im Default wird keine Veränderung der Formanten vorgenommen, bei aktiviertem Button lässt sich ein Wert zwischen 0 und 100 einstellen, der allerdings nicht wie bei Melodyne der Tonhöhe in Cents entspricht.

#### 4.1.2.3 Ableton Live

Eine dritte Software, die der Verfasser zum Verändern der Tonlage verwendete, ist die DAW Ableton Live 9 Suite.

Hier gibt es die Möglichkeit, das

ausgewählte

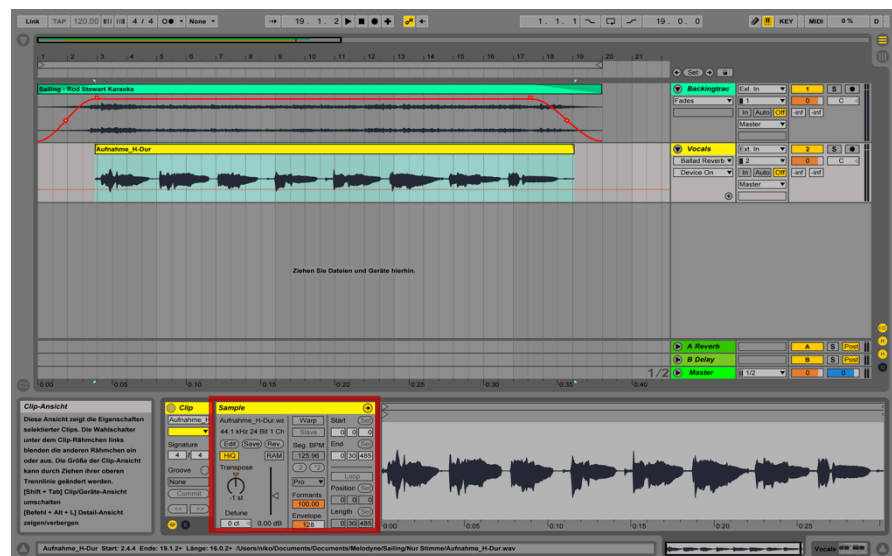


Abbildung 10: Sample-Bearbeitung Ableton Live 9 Suite

Objekt (in diesem Fall den Gesang) direkt im Sample-Fenster zu bearbeiten (vgl. Abb. 10). Genau wie bei Auto-Tune wurde für das Verändern der Tonlage ein Transpose-Rädchen benutzt, welches übrigens eine Transponierung von bis zu vier Oktaven (48 Halbtöne) in jede Richtung zulässt – als Stimme ist das dann allerdings nicht mehr zu erkennen. Auch hier stehen für die Verschiebung unterschiedliche Algorithmen zur Verfügung, die sich aber in diesem Fall nicht auf die Art des Audiosignals beziehen sondern auf die Rechenleistung, die



Ableton bei der Verarbeitung anwendet. Bei diesem Test wurde der rechenaufwändigste Algorithmus Complex Pro gewählt.

Im selben Fenster „Sample“ gibt es ebenfalls ein eigenes Feld zur Bearbeitung der Formanten. Und ebenfalls lässt sich ein Wert zwischen 0 und 100 einstellen, wobei der höchste Wert eine hundertprozentige Erhaltung der Formanten bedeutet und 0 eine Mitverschiebung um den beim Transpose-Rädchen eingestellten Wert.

### **4.1.2.4 Ergebnis**

Wie erwähnt, wurde neben Melodyne und Auto-Tune wurde bei diesem Vergleich auch das Ergebnis der Bearbeitung durch Ableton mit einbezogen.<sup>75</sup> Zur Beurteilung des Ergebnisses wurde erneut auf die Umfrage zurückgegriffen. Diesmal bekamen die Befragten also drei Beispiele zu hören und sollten entscheiden, welches für sie am natürlichsten klingt.<sup>76</sup> Dabei fiel den meisten Befragten die Entscheidung relativ schwer, letztendlich entschieden sich drei Personen für Ableton, vier für Auto-Tune und zwei für Melodyne. Eine der Befragten hörte zwischen Auto-Tune und Melodyne keinen Unterschied, konnte Ableton allerdings ausschließen. Der Verfasser selbst tat sich ebenfalls schwer, sich zwischen Melodyne und Auto-Tune zu entscheiden, entschied sich aber am Ende für Melodyne. Mit einem kleinen Vorsprung wurde also die Verschiebung mit Auto-Tune als das am natürlichsten klingende Beispiel gewählt.

### **4.1.2.5 Exkurs: Identifizierung des Originals**

Aus reinem Interesse von Seiten des Verfassers wurde im Bereich der Tonhöhenverschiebung ein weiteres Experiment durchgeführt, bei dem es nicht um einen direkten Vergleich zwischen Programmen ging, sondern um die Identifizierung der Originalaufnahme aus vier verschiedenen Beispielen, von denen drei in ihrer Tonhöhe verschoben wurden. Als Ausgangsmaterial wurde die gleiche

---

<sup>75</sup> Um ein weiteres DAW-eigenes Plug-In zu testen, wurde auch noch der in Logic Pro X implementierte Pitch Shifter hinzugezogen. Dieser arbeitet allerdings immer mit demselben, nicht veränderbaren Algorithmus, der für die Verarbeitung von Stimme ganz offensichtlich nicht geeignet ist. Das Ergebnis kann dementsprechend in keiner Weise mit den anderen drei Beispielen mithalten und wird nicht weiter beachtet.

<sup>76</sup> Vgl. Hörbeispiele Nr. 16-18.

Aufnahme verwendet wie für den eben beschriebenen Vergleich. Diesmal erfolgte die Verschiebung allerdings auch um einen Halbton nach oben und einen Ganzton nach unten. Für die Verschiebung selbst, allerdings in diesem Zusammenhang nicht wesentlich, wurde Melodyne benutzt.

In der Umfrage wurden dem Teilnehmer nun diese vier Beispiele präsentiert, mit der Aufgabe, das Beispiel herauszufiltern, welches nicht in seiner Tonhöhe verändert, also genauso in dieser Form aufgenommen wurde. Dabei entschied sich eine Teilnehmerin für einen Ganzton unter dem Original, jeweils zwei für einen Halbton darüber und darunter und fünf Personen entschieden sich für das Original. Die Qualität der Bearbeitung war also so gut, dass immerhin die Hälfte der Befragten das Original nicht identifizieren konnte. Außerdem kannten alle Teilnehmenden den Verfasser – und damit Sänger des Materials – persönlich und entsprechend auch seine Stimme. Tatsächlich identifizierte keine der drei Personen, mit denen der Verfasser nur selten in Kontakt steht, die Originalaufnahme.

Der Verfasser selbst kannte natürlich das Original, empfand aber eine Verschiebung um jeweils einen Halbton keinesfalls als unnatürlich. Bei der Verschiebung um einen Ganzton waren allerdings schon etwas deutlicher ein paar digitale Artefakte zu hören. Alle Hörbeispiele finden sich im Anhang.<sup>77</sup>

### **4.1.3 Simulation des „Cher-Effekts“**

1998, ein Jahr nach der Veröffentlichung von Antares Auto-Tune, landete die Sängerin Cher mit dem Titel „Believe“ nach vielen Jahren erneut einen Nummer-Eins-Hit.<sup>78</sup> Das Besondere an dem Song war der roboterartige Effekt auf der Stimme, der im Laufe des Titels immer wieder zu hören war.<sup>79</sup> Der dritte und letzte Teil des Vergleichs bestand darin, diesen auf der Stimme liegenden Effekt mit beiden Anwendungen zu simulieren. Melodyne ist dafür zwar nicht direkt ausgelegt und 1998 wurde der Effekt mit der ersten Version von Auto-Tune erreicht, allerdings ist das Entscheidende für die Entstehung des typischen Klangs die Glättung der Notenlinien und direkte Sprünge, also kein fließender Übergang

---

<sup>77</sup> Vgl. Hörbeispiele Nr. 18-21.

<sup>78</sup> Vgl. Anderson, Trevor (2017): Billboard. Vol. 129, S. 86.

<sup>79</sup> Vgl. Hörbeispiel Nr. 6

zwischen den Noten. Melodyne bietet für beides ein eigenes Werkzeug an, deshalb kann ein Vergleich an dieser Stelle stattfinden. Da auch hier in beiden Programmen mit derselben Aufnahme gearbeitet werden sollte, wurde Auto-Tune auch diesmal erst nach der Aufnahme aktiviert, wie auch schon bei den Tests davor. Ein Exkurs, in dem Auto-Tune live während einer Aufnahme arbeitet, findet sich im weiteren Verlauf dieser Arbeit, kann allerdings für den direkten Vergleich nicht angewendet werden.

### 4.1.3.1 Melodyne

Melodyne bietet wie bereits beschrieben eigene Tools zur Bearbeitung der Notelinie und der Sprünge zwischen den Noten an. Diese können im Noteneditor ausgewählt werden und werden als Tonhöhenmodulations- und Tonhöhendriftwerkzeug bezeichnet. Die Bearbeitung mit diesen Werkzeugen ist ziemlich einfach: Mit dem ausgewählten Werkzeug klickt man auf die zu verändernde Note und bewegt den Zeiger mit gedrückter linker Maustaste so lange nach oben oder unten, bis der gewünschte Effekt erreicht ist. Die Beurteilung erfolgt also zunächst visuell, in Abbildung 11 ist der Unterschied vor und nach der Bearbeitung nochmal verbildlicht.

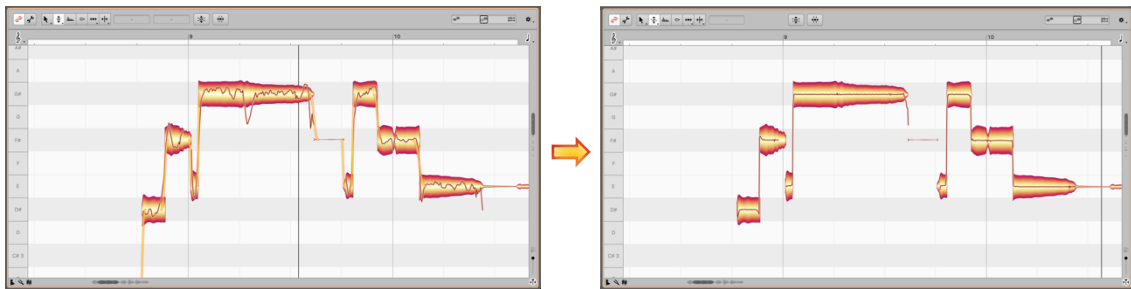


Abbildung 11: Vor und nach der Tonhöhenmodulation

Bei allen Noten, die nur wenig oder gar nicht verändert werden sollen, erfolgt die Bearbeitung entsprechend weniger. Die Bearbeitung ging insgesamt relativ schnell und einfach vonstatten.

### 4.1.3.2 Auto-Tune

1998 wurde der Cher-Effekt mit der ersten Version von Antares Auto-Tune erreicht. Somit lag die Vermutung nahe, man könne einen ähnlich klingenden Effekt

mit einer anderen Stimme mit Hilfe von Auto-Tune Pro leicht erreichen. So einfach wie ursprünglich gedacht war es allerdings nicht.

Der erste Ansatz war, den Retune-Speed auf maximales Tempo einzustellen (in diesem Fall 0, womit das Eingangssignal direkt auf den nächstgelegenen Halbton gepitcht wird), das Vibrato zu deaktivieren und jegliche Parameter, die den Klang natürlicher klingen lassen sollen, auf null zu drehen. Auf diese Weise wurde der Effekt damals erreicht. In diesem Fall war das Ergebnis allerdings ein unnatürlicher Klang, der zwar perfekt intoniert war, aber wenig Ähnlichkeiten mit dem „Cher-Effekt“ aufwies. Auch das Aktivieren des „Classic“ Buttons, was die typischen Klangeigenschaften einer früheren Version verspricht, bewirkte nicht den gewünschten Effekt.

Im Graph Mode gelang es schließlich, dem gewünschten Effekt näher zu kommen. Dafür werden die Notenlinien der zu verändernden Noten „geglättet“, jegliche Modulationen innerhalb einer Note werden also eliminiert, und die Notenlinie wird exakt auf die gewünschte Tonhöhe gezeichnet. Zudem werden alle Noten, die nicht mit dem Effekt belegt werden sollen, per Bypass-Button der Bearbeitung entzogen. Wie bereits in Kapitel 4.1.2.2 kurz beschrieben, ist die Bearbeitung im Graph Mode jedoch vergleichsweise unkomfortabel und etwas unübersichtlich. Trotzdem konnte am Ende ein zufriedenstellendes Ergebnis erstellt werden.

### 4.1.3.3 Ergebnis

Die Beurteilung der Ergebnisse erfolgte auch hier als Teil der Umfrage. Dafür wurde dem Teilnehmer zunächst das Original von Cher vorgespielt, wobei er besonders auf den Effekt auf der Stimme achten sollte, um dann zwischen der Auto-Tune- und der Melodyne-Simulation zu entscheiden, welches Beispiel dem Original näherkommt.<sup>80</sup> Dabei entschieden sich sechs Personen für das Auto-Tune-Beispiel und vier Befragte für Melodyne. Die Mehrzahl empfand also wie der Verfasser selbst das Auto-Tune-Beispiel als näher am Originaleffekt. Damit liegt auch das Programm vorn, mit dem der Effekt ursprünglich erzeugt wurde.

---

<sup>80</sup> Vgl. Hörbeispiele Nr. 7-8.

Trotzdem ist es beachtlich, dass auch Melodyne diesen Klang sehr gut simulieren kann.

## **4.2 Funktionsumfang und Einzelfeatures**

Neben dem gerade beschriebenen direkten Vergleich von Möglichkeiten, die sowohl Melodyne als auch Auto-Tune bieten, soll in diesem Teil auf den jeweiligen Funktionsumfang beider Programme und die entsprechenden Einzelfeatures eingegangen werden, die Melodyne und Auto-Tune in dieser Hinsicht voneinander unterscheiden.

### **4.2.1 Melodyne**

Melodyne bietet einige Funktionen, die in Auto-Tune nicht zu finden sind. Auf drei dieser Features, die den größten Unterschied zu Auto-Tune ausmachen, wird im Folgenden eingegangen.

#### **4.2.1.1 Stand-Alone und Plug-In-Variante**

Ein großer Unterschied, den Melodyne zu Auto-Tune aufweist, ist die Tatsache, dass Melodyne nicht nur als Plug-In, sondern auch als Stand-Alone-Variante funktioniert. In einer Lizenz sind immer beide Varianten enthalten. Dabei macht es keinen Unterschied, welche der vier in Kapitel 3.3 beschriebenen Versionen verwendet wird. Dementsprechend braucht man für die Bearbeitung von Audiomaterial mit Melodyne nicht notwendigerweise eine DAW als Host-Programm. Für den bisherigen Vergleich wurde mit der Plug-In-Variante gearbeitet, im Folgenden wird nun die Stand-Alone-Version beschrieben.

Damit Melodyne überhaupt arbeiten kann, braucht das Programm natürlich ein zu bearbeitendes Audiosignal. Dieses kann auf verschiedene Weise importiert werden: Zum einen lässt sich eine bereits vorhandene Aufnahme oder Audiodatei per Drag and Drop oder über „Ablage/Audio importieren“ einfügen. Zum anderen bietet die Stand-Alone-Variante von Melodyne die sehr komfortable Möglichkeit, eine Aufnahme direkt innerhalb des Programms vorzunehmen. Externe

Hardwaregeräte werden dabei direkt erkannt und sind mit wenigen Klicks als Standardgeräte

festzulegen. Es gibt sogar die Möglichkeit, während der Aufnahme einen vorher importierten Backingtrack und ein Metronom laufen zu lassen. Die

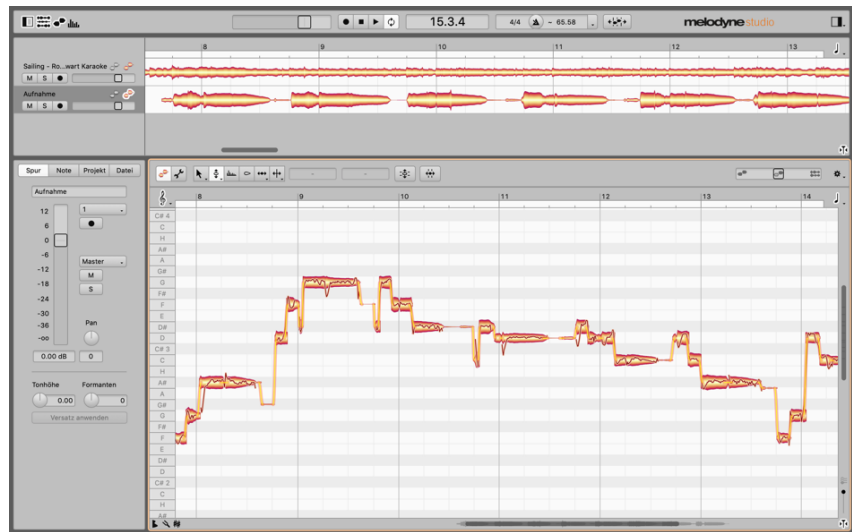


Abbildung 12: Notenerkennung Melodyne

Tempo- und Takterkennung funktionieren dabei problemlos. Allerdings ist das Programm aus Sicht des Verfassers nicht für eine Aufnahme mit Backingtrack gedacht, da auch dieser in seine Notenwerte analysiert und aufgeteilt wird. Nichtsdestoweniger wird die Aufnahme analog zur Plug-In-Version fehlerfrei erkannt und wenige Sekunden nach dem Stoppen der Aufnahme als Notenwerte dargestellt (vgl. Abb. 12).

Außerdem in Abbildung 12 zu erkennen ist die Darstellung jedes Signals als eigene Spur. Diese wird in der Plug-In-Variante durch eine Spurenliste ersetzt, die anzeigt, welcher der Spuren innerhalb der DAW mit einer Melodyne-Instanz belegt sind.

Ein weiterer, aber auch notwendiger Unterschied ist die Möglichkeit des Exportierens eines bearbeiteten Signals. Dabei kann ausgewählt werden, ob jede Spur einzeln oder alle zusammen als Stereomix exportiert werden soll. An dieser Stelle sei nochmal die Möglichkeit erwähnt, das Signal als MIDI-Datei zu exportieren, was eine ganz andere Art der Weiterverarbeitung ermöglicht (vgl. Kap. 3.3).

Die Bearbeitung von Audiomaterial ist in der Stand-Alone-Variante im Vergleich zum Plug-In noch ein Stück komfortabler. Außerdem traten hier keinerlei Performanceeinschränkungen auf. Trotzdem kann es je nach Einsatzgebiet deutlich sinnvoller sein, zum Plug-In zu greifen, beispielsweise für den direkten Vergleich vor und nach einer Korrektur oder im Zusammenhang mit anderen Effekten. In

jedem Fall ist die Möglichkeit, jederzeit zwischen Plug-In- und Stand-Alone-Variante entscheiden zu können, ein nutzerfreundliches Feature von Melodyne.

### 4.2.1.2 Polyphone Bearbeitung

Als Alleinstellungsmerkmal gibt es in Melodyne seit der neuesten Version von 2016 die Möglichkeit, auch polyphones, also mehrstimmiges Material zu bearbeiten. Dies wird durch die neu entwickelte und von Celemony patentierte Technologie DNA (Direct Note Access) ermöglicht (vgl. Kap. 3.3).

Im Rahmen dieser Arbeit wurde nun getestet, wie gut das wirklich funktioniert. Dazu wurde erneut das Kirchenlied „Amazing Grace“ aufgenommen, diesmal allerdings alle Stimmen gleichzeitig auf nur einer Spur. Diese Aufnahme konnte Melodyne weniger gut verarbeiten als angenommen. Zwar wurde jede Stimme im Editor nach der Analyse durch den polyphonen Algorithmus weitgehend vollständig als einzelne Stimme erkannt und angezeigt (vgl. Abb. 13), das Ergebnis der automatischen Korrektur aller Noten

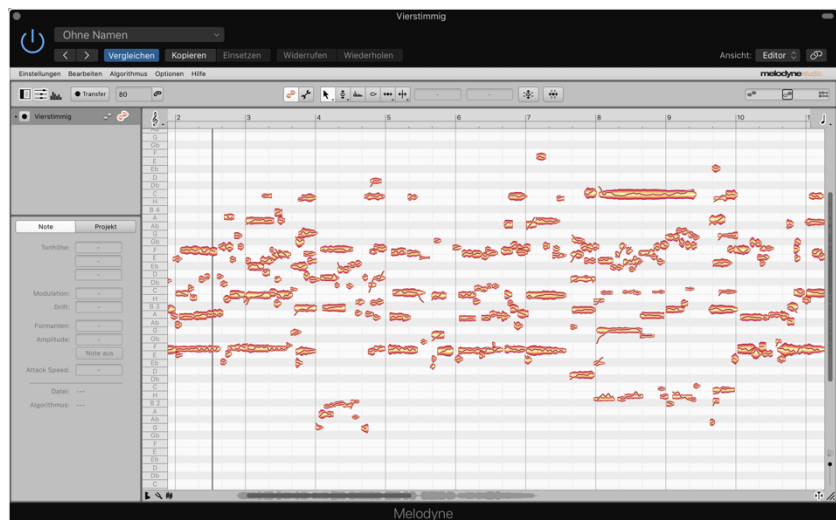


Abbildung 13: mehrstimmige Analyse Melodyne

gleichzeitig fiel allerdings ernüchternd aus. Einzelne Töne wurden in die falsche Richtung gepitcht, es waren sehr viele digitale Artefakte zu hören und der Charakter des Klangs ging verloren. Einzelne falsche Töne individuell und per Hand zu korrigieren, ist natürlich nach wie vor möglich und bringt auch ein gutes Ergebnis, dies ist jedoch mit sehr viel Arbeit und einem ständigen Abgleichen mit dem Arrangement verbunden. Außerdem muss man dabei darauf achten, keine Obertöne zu verändern, die durch die Analyse teilweise als eigene Notenwerte angezeigt werden.

Testweise wurde auch eine zweistimmige Version aus den einzelnen Stimmen gebastelt und Melodyne zur Analyse angeboten. Damit hatte Melodyne schon deutlich weniger Probleme, bei der automatischen Korrektur wurden allerdings nach wie vor einzelne Töne in die falsche Richtung korrigiert und die Obertonstruktur teilweise verändert.

#### **4.2.1.3 Exkurs: Duplizieren von Stimmen**

Als ein weiteres Experiment im Bereich der Mehrstimmigkeit wurde versucht, anhand von Melodyne jeweils eine Stimme durch Duplizieren und extreme Bearbeitung eine andere Stimme zu ersetzen. Als Testmaterial wurden die Einzelstimmen der Amazing-Grace-Aufnahme verwendet. So wurde beispielsweise die Bassstimme dupliziert und so verändert, dass sie den Noten des Tenors entsprach. Das Gleiche passierte mit der duplizierten Spur der Alt-Stimme, womit auch die Sopran-Stimme abgedeckt war. Die Bearbeitung in Melodyne war dabei denkbar einfach. Grundsätzlich hat sich herausgestellt, dass das Höherpitchen von tieferen Stimmen besser funktioniert als andersherum. Trotzdem klingt das Ergebnis sehr unnatürlich<sup>81</sup> und kann nach Meinung des Verfassers in der Praxis nur in absoluten Extremfällen bei beispielsweise Backingvocals angewendet werden, bei denen gleichzeitig so viel anderes passiert, dass der Effekt nicht mehr hörbar ist.

#### **4.2.1.4 Sound Editor**

Als drittes und letztes Einzelfeature von Melodyne wird hier der Sound Editor genannt. Damit lässt sich der Klangcharakter einer Aufnahme auf verschiedene Weise verändern. Zum einen gibt es einen eigenen Equalizer, womit sich einzelne Frequenzen verstärken oder abschwächen lassen bzw. das gesamte Frequenzspektrum verändert werden kann. Zum anderen bietet der Sound Editor sehr umfangreiche Möglichkeiten an, die Obertonstruktur und den Charakter eines Klanges zu bearbeiten. Beispielsweise lässt sich jeder harmonische Oberton bis zu etwa 70. Stelle individuell anpassen, verschiedene Muster sind einstellbar,

---

<sup>81</sup> Vgl. Hörbeispiel Nr. 15.



sodass z.B. jeder zweite Oberton gar nicht mehr oder nur auf sehr geringer Lautstärke hörbar ist usw. Außerdem kann ein Klang nach den eigenen Vorstellungen bezüglich Frequenzspektrum, Formanten, Amplitude und Wellenform neu erzeugt werden („resynthesize“). Damit wäre beispielsweise möglich, die Obertonstruktur einer Violine auf die Stimme anzuwenden oder andersherum. Spätestens hier wird eine Bearbeitung allerdings sehr komplex und an dieser Stelle nicht weiter vertieft.

### 4.2.2 Auto-Tune

Die einzige wirklich relevante Zusatzfunktion, die Auto-Tune im Vergleich zu Melodyne besitzt, stellt gleichzeitig die Hauptarbeitsweise des Programms dar: die Korrektur und Veränderung des eingehenden Signals in Echtzeit. Hier ist dennoch unbedingt zwischen Echtzeit und Live zu unterscheiden. Echtzeit bedeutet, dass ein bereits aufgenommenes Signal zur Bearbeitung an Auto-Tune geschickt wird und während des Abspielens korrigiert bzw. verändert wird. Live hingegen bedeutet, dass eine Bearbeitung entweder während der Aufnahme oder eben live auf der Bühne bei einem Konzert vonstatten geht. Alle in Kapitel 4.1 beschriebenen Tests wurden zugunsten der Vergleichbarkeit wie beschrieben *nach* der Aufnahme durchgeführt. Da die Live-Funktion von Auto-Tune allerdings einen großen Anwendungsbereich des Programms darstellt, soll diese im Folgenden nochmal gesondert aufgeführt werden.

In diesem Zusammenhang wurde nicht mit Auto-Tune Pro, sondern mit Auto-Tune Live gearbeitet, da diese Version von Auto-Tune mit möglichst geringer Latenz arbeitet und damit speziell für den Live-Bereich konzipiert ist. Das Programm leistet dabei sehr gute Arbeit.

Die Funktionsweise und einstellbaren Parameter von Auto-Tune Live wurden bereits im letzten Kapitel (vgl. Kap. 3.2.2) ausführlich beschrieben und deshalb an dieser Stelle nicht weiter vertieft.

Getestet wurden verschiedene Einstellungen von Auto-Tune Live, wobei diese für jede neue Aufnahme verändert wurden. Eine Aufnahme erfolgte mit dem Ziel, falsche Töne zu korrigieren, die Natürlichkeit und den Klangcharakter dabei allerdings zu erhalten. Bei der zweiten Aufnahme wurden der Retune-Speed, das

Vibrato und der Humanize-Parameter absichtlich auf null gedreht, um den „Cher-Effekt“ im Live-Einsatz zu erzeugen. Jeweils ein Beispiel ohne Auto-Tune-Einsatz, eins mit „natürlicher“ Korrektur und eins mit dem „Cher-Effekt“ finden sich im Anhang.<sup>82</sup> In diesem Zusammenhang fiel dem Verfasser des Weiteren auf, dass der „Cher-Effekt“ mit Auto-Tune Live anhand der Parameter erreicht wurde, die denselben Effekt vorher mit Auto-Tune Pro in der Nachbearbeitung des bereits aufgenommenen Materials nicht erreichen konnten. Dieser gelang also mit Auto-Tune Live während der Aufnahme mit deutlich weniger Aufwand als mit Auto-Tune Pro im Nachhinein. Daraus lässt sich schließen, dass Auto-Tune im Live-Bereich anders arbeitet als bei der Echtzeitbearbeitung von bereits aufgenommenem Material.

### **4.3 Benutzeroberfläche und Handhabung**

Generell kann man sagen, dass Melodyne und Auto-Tune auf völlig unterschiedlichen Herangehensweisen beruhen. Auto-Tune (abgesehen vom Graph Mode) fängt an zu arbeiten, sobald eine Instanz in den Kanalzug der Spur geladen wurde. Es pitcht jedes eingehende Signal je nach den eingestellten Parametern in eine bestimmte Richtung und arbeitet dabei in Echtzeit. Melodyne auf der anderen Seite muss, nach dem Einspielen des Materials, das Signal erstmal analysieren und nimmt ohne Bearbeitung des Anwenders zunächst keinerlei Veränderungen vor.

Auch die Vorgehensweise bei der Bedienung an sich unterscheidet sich stark: Während Melodyne auch eine große visuelle Komponente bei der Bearbeitung aufweist, ist das Arbeiten mit Auto-Tune als „technischer“ zu beschreiben.

Die Benutzeroberfläche und Handhabung beider Programme werden im Folgenden etwas genauer beschrieben.

---

<sup>82</sup> Vgl. Hörbeispiele Nr. 9-11.

### **4.3.1 Melodyne**

Zunächst wird die Benutzeroberfläche betrachtet. Diese ist aus Sicht des Verfassers sehr übersichtlich und man findet sich schnell zurecht. Jedes Fenster hat seinen Platz und lässt sich zum einen ein- und ausschalten, zum anderen in seiner Größe verändern, wenn man für die Bearbeitung oder Übersicht in einem anderen Fenster mehr Platz braucht.

Besonders gelungen ist das Noteneditorfenster. Jedes analysierte Signal wird in seinen einzelnen Noten auf einer Klaviatur dargestellt, wobei jede Note genau dort positioniert ist, wo sie sich intonationsmäßig befindet. Wenn also beispielsweise eine Note um 25 Cents höher als beabsichtigt gesungen wurde, ist das im Noteneditor auch so dargestellt. Dadurch wird bei der Bearbeitung auch visuell sehr schön sichtbar, wie gut eine Note intoniert ist und inwiefern Veränderungen vorgenommen werden.

Die Bedienung ist intuitiv und man entwickelt ziemlich schnell einen eigenen Workflow, wenn man sich ein paar Stunden an die Bearbeitung setzt. Melodyne bietet dabei einige Hilfen, die ein schnelles Arbeiten zulassen. Dazu gehören zum Beispiel das Setzen einer falschen Note auf den nächstgelegenen Halbton durch Doppelklick oder der einfache Wechsel zwischen verschiedenen Arbeitswerkzeugen durch Rechtsklick innerhalb des Noteneditors.

Anhand des Infofensters ist jederzeit erkennbar, welche Bearbeitungen im Bereich des ausgewählten Reiters (Spur, Note, Projekt, Datei) getätigt wurden bzw. welche Dateien verwendet wurden und wo diese zu finden sind.

### **4.3.2 Auto-Tune**

Da im Verlauf zwar zum größten Teil mit Auto-Tune Pro gearbeitet wurde, aber auch Auto-Tune Live zum Einsatz kam und sich deren Benutzeroberflächen unterscheiden, wird auf beide Versionen kurz eingegangen.

#### **4.3.2.1 Auto-Tune Pro**

In puncto Übersichtlichkeit ist auch die Oberfläche des Auto Mode von Auto-Tune Pro sehr gelungen. Alle wichtigen Parameter sind dargestellt und mit wenig Aufwand nach Belieben einzustellen. Für eine genauere Bearbeitung der Tonart (beispielsweise die Bypass-Funktion) ist die Möglichkeit der De- und Aktivierung der Advanced-Ansicht eine sehr elegante Lösung.

Der Graph Mode auf der anderen Seite ist eher unübersichtlich. Hier braucht man eine Weile, bis man sich zurechtgefunden hat. Für eine Bearbeitung innerhalb des Graph Modes muss das Signal zunächst ähnlich wie bei Melodyne transferiert werden. Im Anschluss daran wird das Signal allerdings in seiner Wellenform und nicht etwa als einzelne Notenwerte dargestellt. Dies passiert erst anhand von Linien, die eingezeichnet werden, nachdem der „Make Curves“ Button betätigt wurde. Durch die farbige Abgrenzung sollen die Unterschiede zwischen dem tatsächlichen, dem Ziel- und dem resultierenden Notenwert verbildlicht werden, allerdings liegen diese teilweise sehr nahe beieinander, wodurch ein eher chaotisches Bild entsteht. Das wurde in Melodyne sehr viel besser gelöst.

#### **4.3.2.2 Auto-Tune Live**

Das Übersichtsfenster von Auto-Tune Live hingegen wirkt zwar auf den ersten Blick etwas unsortiert, ergibt aber durchaus Sinn, sobald man eine Weile damit gearbeitet hat. Man könnte diese Ansicht als „technischer“ beschreiben. Hier ist der „Advanced Mode“ quasi in die Standardansicht mit eingepflegt, der gleichzeitig auch das Zentrum des Fensters darstellt. Ein kleiner Vorteil davon kann natürlich sein, dass auf jeden Parameter direkt zugegriffen werden kann und sich keine Einstellungen in einem Advanced oder Graph Mode „verstecken“.

### **4.4 Zwischenfazit**

Sowohl Melodyne als auch Auto-Tune erfüllen sehr zufriedenstellend ihren Zweck. Beide sind für die Tonhöhenkorrektur von Audiomaterial konzipiert, arbeiten allerdings unterschiedlich. Ihr größter Unterschied liegt darin, dass Auto-Tune eher für eine Echtzeitkorrektur und Melodyne ausschließlich für die

Nachbearbeitung entwickelt wurde. Daraus ergeben sich die jeweiligen Stärken und Schwächen der Programme: Auto-Tune bietet die Möglichkeit, schon während der Aufnahme bzw. in Echtzeit gute Ergebnisse zu liefern, wobei allerdings die vorherige Einstellung aller Parameter zeitaufwendig sein kann. Mit Melodyne lässt sich eine Aufnahme im Nachhinein sehr schnell und sehr gut korrigieren. Das Programm bietet außerdem die Möglichkeit, sehr stark in den Charakter und die Struktur eines Klanges einzugreifen und diesen völlig unkenntlich zu machen oder komplett neu zu gestalten. Des Weiteren ist die Möglichkeit zur Bearbeitung von polyphonem Material eine aus Praxissicht grundsätzlich begrüßenswerte Neuerung im Bereich der Tonhöhenkorrektur und -manipulation. Allerdings muss diese Funktion noch deutlich verbessert werden, da bei der Analyse und Verarbeitung von vierstimmigem Gesangsmaterial auch Melodyne erkennbar an seine Grenzen kommt.

Eine Veränderung der Tonart ist für beide Programme kein Problem, allerdings wird die Bearbeitung immer deutlicher hörbar, je weiter man sich von der Originaltonart wegbewegt.<sup>83</sup>

Im direkten Klangvergleich zwischen Auto-Tune und Melodyne in verschiedenen Bereichen war kein eindeutiger Sieger festzustellen, wobei Auto-Tune insgesamt laut Umfrage etwas besser abgeschnitten hat.

Der Verfasser selbst war mit beiden Programmen sehr zufrieden, wobei ihm vor allem die Arbeitsweise und das Gesamtpaket bei Melodyne mehr zugesagt hat.

Abschließend lässt sich sagen, dass Antares mit Auto-Tune im Live-Bereich ohne Frage vergeblich seinesgleichen sucht, man aber für professionelle Nachbearbeitung im Studiobereich mit Melodyne einerseits insgesamt mehr Möglichkeiten geboten bekommt und eine Bearbeitung durch die zusätzliche visuelle Komponente deutlich intuitiver erfolgen kann.

---

<sup>83</sup> An dieser Stelle sei auch kurz nochmal der Transpose-Algorithmus von Ableton Live 9 Suite erwähnt, der ebenfalls gute Arbeit leistet.

## 5 Anwendungsbereiche

Tonhöhenkorrektur findet in verschiedenen musikalischen Bereichen Anwendung. Daniel Griffiths, Chefredakteur der Musikzeitschrift „Future Music“ sagte bereits 2010, dass Tonhöhenkorrektur in 99% der im kommerziellen Bereich aufgenommenen Musik genutzt wird.<sup>84</sup> Abgesehen davon, dass in nahezu jeder Pop-Produktion unhörbare Tonhöhenkorrektur stattfindet, wird der Auto-Tune-Effekt auch immer wieder als Stilmittel eingesetzt. Im Folgenden wird auf diese Beispiele genauer eingegangen.

### 5.1 Perfekte Intonation

Der am wenigsten hörbare aber mit Abstand am weitesten verbreitete Einsatz von Plug-Ins wie Auto-Tune oder Melodyne findet sich in nahezu jeder modernen Pop-Produktion. Hier werden kleine Korrekturen in der Tonhöhe gemacht, die keinesfalls zu hören sein sollen, das Ganze allerdings noch besser klingen lassen. Seien es Harmonien der Backingvocals oder ein saubererer Übergang zwischen zwei Tönen, die gesamte Produktion wird bis zur Perfektion geglättet.<sup>85</sup> Drew Waters, Geschäftsführer und Produzent des Studios Capitol Records sagte:

*„I’ll be in a studio and hear a singer down the hall and she’s clearly out of tune, and she’ll do one take.“<sup>86</sup>*

Und so ähnlich sei das auch bei vielen bekannten Künstlern der Fall. Madonna, Lady Gaga, Maroon 5, Ke\$ha und viele, viele weitere gehören dazu. Die wenigsten allerdings geben es zu.<sup>87</sup> Dass Tonhöhenkorrektur ist den meisten Produktionen verwendet wird, ist eher als „Common Knowledge“ zu verstehen, über die man als Künstler aber nur in den seltensten Fällen spricht. Eine Ausnahme macht

---

<sup>84</sup> Vgl. Savage, Mark (2010): How commonplace is auto-tune? <https://www.bbc.co.uk/news/entertainment-arts-11056840> (Zugriff: 08.02.2019).

<sup>85</sup> Vgl. Anderson, Lessley (2013): Seduced by ‘perfect’ pitch: How Auto-Tune conquered pop music, <https://www.theverge.com/2013/2/27/3964406/seduced-by-perfect-pitch-how-auto-tune-conquered-pop-music> (Zugriff: 27.01.2019).

<sup>86</sup> Vgl. Waters, Drew, zitiert nach Anderson, Lessley.

<sup>87</sup> Vgl. Anderson, Lessley (2013): Seduced by ‘perfect’ pitch: How Auto-Tune conquered pop music, Link vgl. oben, (Zugriff: 27.01.2019).

dabei Justin Bieber. Der kanadische Sänger gab 2012 gegenüber dem Q magazine zu, dass seine Produzenten für sein drittes Album „Believe“ seine Stimme anhand von Melodyne korrigiert hätten. Weiterhin sagte er, 99% aller Künstler täten dies und kein Mensch treffe jede Note perfekt.<sup>88</sup>

Als einen der größten Faktoren, warum eine korrekte Intonation so wichtig sei, nennt Andy Hildebrand, der Erfinder von Auto-Tune, die Emotionalität:

*„Voices or instruments are out of tune when their pitch is not sufficiently close to standard pitches expected by the listener, given the harmonic fabric of the example. When voices or instruments are out of tune, the emotional qualities of the performance are lost. Correcting intonation [...] solves this problem and restores the performance.“<sup>89</sup>*

Ohne korrekte Intonation könne also die Emotionalität nicht übertragen werden, die die eigentliche „Message“ beinhalte. Tonhöhenkorrektur sei demnach also nur das Medium zur Übertragung ebendieser Message.

## 5.2 Zeitersparnis

Abgesehen von der perfekten Intonation ist die Benutzung von Auto-Tune oder Melodyne aber auch ein Zeitfaktor im Studio. Ein Aspekt, der bei Diskussionen oft auf der Strecke bleibt. Ebenfalls Andy Hildebrand sagte darüber:

*„[Auto-Tune’s] largest effect in the community is it’s changed the economics of sound studios. Before Auto-Tune, sound studios would spend a lot of time with singers, getting them on pitch and getting a good emotional performance. Now they just do the emotional performance, they don’t worry about the pitch, the singer goes home, and they fix it in the mix.“<sup>90</sup>*

Auto-Tune wird also oft auch schon während einer Aufnahme verwendet, um sich den Aufwand der nachträglichen Korrektur zu sparen.

---

<sup>88</sup> Vgl. The Daily Telegraph (2012): Sydney bound Justin Bieber admits voice tuned by software Melodyne for third album Believe, <https://www.dailytelegraph.com.au/entertainment/sydney-bound-justin-bieber-admits-voice-tuned-by-software-melodyne-for-third-album-believe/news-story/43672f02f2103f812990e1232f1b36f5> (Zugriff: 09.02.2019).

<sup>89</sup> Vgl. Hildebrand, Andy, zitiert nach Marshall, Owen (2014): A Brief History of Auto-Tune, <https://soundstudies-blog.com/2014/04/21/its-about-time-auto-tune/> (Zugriff: 08.02.2019).

<sup>90</sup> Vgl. ebd. (Zugriff: 09.02.2019).

Und auch außerhalb vom Tonstudio findet Auto-Tune im Live-Bereich Verwendung. Ein Sänger oder eine Sängerin kann also während eines Konzertes sehr schlecht intonieren, die Fehler werden allerdings korrigiert, bevor sie das Publikum erreichen.<sup>91</sup> Sogar bei der 7. Staffel der 2010 in Großbritannien ausgestrahlten Casting-Show „X Factor“, bei der es darum geht, dass die Kandidaten allein durch ihre Stimmen überzeugen müssen, wurde mit Auto-Tune getrickst, um die Show für die Zuschauer unterhaltsamer zu machen.<sup>92</sup>

Abgesehen von der perfekten Intonation führen extreme Einstellungen, insbesondere bei Auto-Tune zu einem roboterartigen Effekt auf der Stimme. Wie sehr dieser bereits mehrfach erwähnte „Cher-Effekt“ in der Praxis eingesetzt wird, damit befasst sich der folgende Abschnitt.

### 5.3 „Cher-Effekt“

Erreicht wurde der Cher-Effekt, indem man die Retune Speed, also die Zeit, die Auto-Tune hat, um das Eingangssignal zu pitchen, auf den kleinstmöglichen Wert festlegt (vgl. Kap. 3.2.1). Dabei wird gleichzeitig jede Form von Variation innerhalb der Note eliminiert und die Sprünge zwischen zwei Noten erfolgen augenblicklich und ohne Übergang. Der typische Klang des Effekts ist dabei sehr bekannt und wurde schon an mehreren Beispielen im Verlauf dieser Arbeit aufgezeigt.

Der Cher-Effekt war allerdings nicht der erste roboterartige Sound seiner Art. Vergleichbare Effekte wurden bereits in den 1970er Jahren anhand von verschiedener Hardware wie dem Vocoder bzw. der Talkbox erreicht. Bei der Talkbox wird ein elektronisches Signal in Schall umgewandelt und über einen Schlauch in den Mundraum des Anwenders geleitet. Der Resonanzraum wird also durch Stellung von Zunge und Lippen verändert. Der Vocoder hingegen analysiert das Eingangssignal und filtert dieses anhand von elektronischen Filtern, die den

---

<sup>91</sup> Vgl. Anderson, Lessley (2013): Seduced by ‘perfect’ pitch: How Auto-Tune conquered pop music, <https://www.theverge.com/2013/2/27/3964406/seduced-by-perfect-pitch-how-auto-tune-conquered-pop-music> (Zugriff: 09.02.2019).

<sup>92</sup> Vgl. BBC News (2010): X Factor admits tweaking vocals, <https://www.bbc.com/news/entertainment-arts-11056050> (Zugriff: 09.02.2019).



Vokaltrakt und damit die Formanten nachbilden. Mit Auto-Tune allerdings wurde dieser Effekt als Software für ein deutlich breiteres Publikum zugänglich.

In den folgenden Jahren verbreitete sich dieser charakteristische Effekt in der Branche und wird bis heute vor allem in den Bereichen Hip-Hop, R&B und Dance Music vielfach eingesetzt.<sup>93</sup> Unter anderen sind die Rapper Kanye West mit seinem Album „808s & Heartbreak“ und T-Pain mit nahezu jedem seiner Songs sehr bekannt für den Einsatz dieses Effektes.<sup>94</sup> Seit 2009 hat sich in den USA sogar ein eigenes Subgenre des Hip-Hop entwickelt, Cloudrap genannt, welches sich vor allem durch sphärische Klänge und die intensive Nutzung von Auto-Tune auszeichnet. Ein Songbeispiel eines bekannten US-amerikanischen Vertreters findet sich im Anhang.<sup>95</sup>

Andy Hildebrand sagte in einem Interview, in dem er nach seiner Meinung zur Nutzung seines Programms als ebendiesen Effekt gefragt wurde:

*„I just make the car. I don't drive it down the wrong side of the road.“<sup>96</sup>*

Er konzipierte Auto-Tune also ursprünglich für die Korrektur falscher Noten. Die Chers und T-Pains dieser Welt sind nach dieser Analogie verrückte Geisterfahrer.

---

<sup>93</sup> Vgl. Anderson, Lessley (2013): Seduced by `perfect´ pitch: How Auto-Tune conquered pop music, <https://www.theverge.com/2013/2/27/3964406/seduced-by-perfect-pitch-how-auto-tune-conquered-pop-music> (Zugriff: 27.01.2019).

<sup>94</sup> Vgl. Hörbeispiel Nr. 4.

<sup>95</sup> Vgl. Hörbeispiel Nr. 5.

<sup>96</sup> Vgl. Hildebrand, Andy, zitiert nach Anderson, Lessley (2013): Seduced by `perfect´ pitch: How Auto-Tune conquered pop music, Link vgl. oben (Zugriff: 10.02.2019).

## 6 Kritik und Entwicklung

Wie es häufig der Fall mit neuer Musiksoftware ist, waren auch Auto-Tune und Melodyne Gegenstand von Kontroversen und Kritik und sind es bis heute. Inwiefern sich diese äußert und welche Entwicklungen dieser Bereich der Audiomanipulation erlebt, wird im Folgenden kurz beschrieben.

### 6.1 Kritik

Seit ihrer Erfindung sind Auto-Tune & Co immer wieder negativer Kritik ausgesetzt. In einem Internetforum wurde gepostet, derjenige, der Auto-Tune erfunden hat, müsse Musik hassen. Man sprach von Entwertung des tatsächlichen Talents.<sup>97</sup>

Larry Crane, Herausgeber der Musikzeitschrift „Tape Op“ und Musikproduzent argumentierte, Auto-Tune sei im Gegensatz zu anderen Effekten unnatürlich:

*„The thing is, reverb and delay always existed in the real world, by placing an artist in unique environment, so [those effects are] just mimicking reality. [...] There is nothing in the natural world that Auto-Tune is mimicking, therefore any use of it should be carefully considered.“<sup>98</sup>*

Als Gegenargument könnte angeführt werden, dass auch die Benutzung eines Kompressors beispielsweise in diese Kategorie fällt.

Aber auch Künstler äußerten sich negativ gegenüber der Verwendung von Programmen wie Auto-Tune und Melodyne. Die US-amerikanische Country-Sängerin Allison Moorer beispielsweise versah ihr Album „Miss Fortune“ (2002) deutlich sichtbar mit dem Hinweis, dass keinerlei Stimmanpassung oder Tonhöhenkorrektur beim Aufnehmen dieses Albums verwendet wurde. Der US-amerikanische Rapper Jay-Z veröffentlichte 2009 einen Titel mit dem Namen Death of Auto-Tune (D.O.A), in dem er seinen Unmut gegenüber dessen Benutzung ausdrückt.

---

<sup>97</sup> Vgl. Anderson, Lessley (2013): Seduced by 'perfect' pitch: How Auto-Tune conquered pop music, <https://www.theverge.com/2013/2/27/3964406/seduced-by-perfect-pitch-how-auto-tune-conquered-pop-music> (Zugriff: 09.02.2019).

<sup>98</sup> Vgl. Crane, Larry, zitiert nach Andersson, Lessley (2013): Seduced by 'perfect' pitch: How Auto-Tune conquered pop music, Link vgl. oben (Zugriff: 10.02.2019).

Auto-Tune Erfinder Andy Hildebrand selbst wurde in verschiedenen Interviews schon mehrfach gefragt, ob Auto-Tune böse sei bzw. sinngemäße ähnliche Formulierungen. Auf solche Art von Fragen stelle er meist relativ trocken eine Gegenfrage:

*„My wife wears makeup, does that make her evil?“<sup>99</sup>*

## 6.2 Entwicklung

Unabhängig davon, dass sich Programm zur Tonhöhenmanipulation wie Melodyne und Auto-Tune in den nächsten Jahren stetig weiter entwickeln werden, befinden sich bereits seit einigen Jahren verschiedene Softwares in der Entwicklung, die darüber hinausgehen und mit Stimmsynthese im direkten Zusammenhang mit Text experimentieren. Eines dieser Programme heißt Vocaloid und soll im Folgenden stellvertretend als Beispiel vorgestellt werden.

Die Software **Vocaloid** der Firma Yamaha ist ein Synthesizer, der mittels Sprachsynthese künstlichen Gesang erzeugen kann.<sup>100</sup> Das Programm wurde 2003 vorgestellt und ist mittlerweile bei der Version „VOCALOID5“ angekommen. Es ist vor allem in Japan und den USA verbreitet.

Die Funktionsweise lässt sich wie folgt beschreiben: Je nach Versionsumfang (Standard und Premium) gibt es vier bzw. acht sogenannte Voicebanks, also Stimmen, auf denen die weitere Bearbeitung beruht. Die Standardversion bietet jeweils eine männliche und eine weibliche Stimme auf Englisch und Japanisch an. Nun gibt es zwei Möglichkeiten, ein eigenes Sample zu erstellen. Einerseits kann eines von über 1000 von VOCALOID bereit gestellten Samples so verändert werden, dass es den eigenen Vorstellungen entspricht. Dazu wählt man Stimme und Sprache und entscheidet sich für ein oder mehrere verschiedene Attribute (z.B. Bright, Dreamy, Shout usw.) die das Sample aufweisen soll. Daraus ergibt sich dann eine Liste von „vocal phrases“, also Samples, die diesen gewählten Parametern entsprechen. Davon wird nun eines gewählt und zur

---

<sup>99</sup> Vgl. Anderson, Lessley (2013): Seduced by `perfect` pitch: How Auto-Tune conquered pop music, <https://www.theverge.com/2013/2/27/3964406/seduced-by-perfect-pitch-how-auto-tune-conquered-pop-music> (Zugriff: 13.02.2019).

<sup>100</sup> Vgl. VOCALOID (2018): VOCALOID5 – Walkthrough, <http://www.vocaloid.com/en/> (Zugriff: 13.02.2019).

Bearbeitung in eine Editorfenster bewegt. Per Doppelklick auf das Sample wird dieses nun in einer Art Piano Roll in seinen einzelnen Notenwerten angezeigt. Dabei ist jede Silbe des Gesangs der entsprechenden Note zugeordnet. Nun können sowohl die einzelnen Notenwerte als auch der Text bearbeitet werden. Zusätzlich lässt sich jede Note individuell mit verschiedenen Arten von Vibrato versehen oder die Stimme an sich in ihrer Farbe bearbeiten. Auch die Belegung des gesamten Samples mit gängigen Audioeffekten wie Hall, Kompressor usw. ist möglich.

Andererseits lässt sich auch ein komplett neues Sample von null an erstellen. Sind Sänger und Sprache ausgewählt, wird dazu auf eine Spur eine zunächst leere Datei erzeugt, die dann bearbeitet werden kann. Die Noten werden in diesem Fall per Hand eingezeichnet und sprachlich zunächst mit der Silbe „Ouh“ belegt. Dieser Text kann nun angepasst und einer gewünschten Note zugeordnet werden. Auch hier können die oben beschriebenen Effekte wie Vibrato usw. angewendet werden.<sup>101</sup>

Mit Programmen wie diesem lassen sich also Gesangspassagen in allen möglichen Formen und Tonarten kreieren und dementsprechend auch Harmonien erzeugen. Theoretisch ist es demnach denkbar, beispielsweise einen kompletten Chor zu simulieren, ohne selbst jemals eine eigene Aufnahme gemacht zu haben. Eine Originalaufnahme klingt momentan immer noch weitaus natürlicher, allerdings werden sich diese Unterschiede in den folgenden Jahren immer weiter zurückbilden, bis ein Unterschied zwischen dem Original und einer digitalen Simulation kaum noch zu hören sein wird.

---

<sup>101</sup> Vgl. VOCALOID (2018): VOCALOID5 – Walkthrough, <http://www.vocaloid.com/en/> (Zugriff: 13.02.2019).

## 7 Fazit

Nach der Betrachtung der beiden großen Programme im Bereich der Tonhöhenkorrektur wird deutlich, welche technischen Möglichkeiten bereits existieren und was in der Zukunft möglich sein wird.

Durch den Einsatz von Tonhöhenkorrektur kann eine Intonation geschaffen werden, die über die natürlichen Möglichkeiten eines Menschen hinausgeht. Dies ist vergleichbar mit einem Foto in einem Beauty-Magazin. Auch hier wird das Foto eines Models bis hin zur übernatürlichen Perfektion bearbeitet, wodurch eine falsche Realität geschaffen wird. Genauso lässt sich perfekt intonierter Gesang auf Sänger und Sängerinnen übertragen, die anhand von technischen Mitteln idealisiert werden. In diesem Zusammenhang stellt sich unweigerlich auch die Frage nach der Authentizität. Insbesondere, wenn eine Korrektur weniger hörbar ist, kann dies auch als Kaschierung mangelnder stimmlicher Fähigkeiten interpretiert werden. Wie wichtig sind diese demnach heutzutage noch für die Bekanntheit und vor allem Bewertung eines Künstlers, insbesondere im Vergleich zum äußeren Erscheinungsbild, Medienpräsenz usw.?

Auf der anderen Seite bietet solche Software auch Möglichkeiten zur kreativen Nutzung. Es lassen sich Harmonien erschaffen, der Klangcharakter einer Stimme kann verändert und Effekte verschiedener Art erzeugt werden.

Fest steht, die technischen Möglichkeiten sind in jedem Fall gegeben und entwickeln sich stetig weiter. Und diese technischen Möglichkeiten nach heutigem Stand zu testen und untereinander zu vergleichen, war Ziel dieser Arbeit.

Zu welchem Programm man nun am ehesten greifen sollte, hängt immer davon ab, in welchem Gebiet es eingesetzt werden soll. Das eine Programm für alle Anwendungen gibt es noch nicht. Für den Live-Bereich ist sicherlich Auto-Tune von Antares die beste Wahl. Melodyne von Celemony hingegen ist aufgrund der einfachen Bedienung und des großen Funktionsumfangs aus Sicht des Verfassers für den Studioalltag besser geeignet.

---

## Literaturverzeichnis

### Monographien

Dickreiter, Michael (2011): Mikrofonaufnahme, 4. Auflage, Stuttgart.

Friesecke, Andreas (2013): Studio Akustik, 4. Auflage, Bergkirchen.

Große, Günter (1981): Von der Edisonwalze zur Stereoplatte. 1. Auflage, Berlin.

Katz, Mark (2010): Capturing Sound. How technology has changed music, London.

Manning, Peter (1993): Electronic & Computer Music, 2. Auflage, USA.

Roads, Curtis (2004): Microsound, USA.

Roads, Curtis (1996): The computer music tutorial, USA.

Watkinson, John (1994): The Art of Digital Audio, 2. Auflage, GB.

Zölzer, Udo (1996): Digitale Audiosignalverarbeitung, Stuttgart.

### Sammelbände

Cipriani, Alessandro; Giri, Maurizio (2014): Electronic Music and Sound Design. 1. Auflage, Rom. Übersetzung ins Englische von Richard Dudas.

Dickreiter, Michael; Dittel, Volker; Hoeg, Wolfgang; Wöhr, Martin (2008): Handbuch der Tonstudioteknik, Band 1. 7. Auflage, München.

Dickreiter, Michael; Dittel, Volker; Hoeg, Wolfgang; Wöhr, Martin (2014): Handbuch der Tonstudioteknik, Band 2, 8. Auflage, Berlin.

## Bedienungsanleitungen

Antares Audio Technologies (2016): Auto-Tune EFX 3, Owner's Manual. URL: [https://www.atarestech.com/mediafiles/documentation\\_records/11\\_Auto-Tune\\_EFX\\_3\\_Manual.pdf](https://www.atarestech.com/mediafiles/documentation_records/11_Auto-Tune_EFX_3_Manual.pdf) (Zugriff: 18.01.2019).

Antares Audio Technologies (2016): Auto-Tune Live. Owner's Manual. URL: [https://www.atarestech.com/mediafiles/documentation\\_records/10\\_Auto-Tune\\_Live\\_Manual.pdf](https://www.atarestech.com/mediafiles/documentation_records/10_Auto-Tune_Live_Manual.pdf) (Zugriff: 16.01.2019).

Antares Audio Technologies (2018): Auto-Tune Pro. User Guide. URL: [https://www.atarestech.com/mediafiles/documentation\\_records/Auto-Tune\\_Pro\\_Manual.pdf](https://www.atarestech.com/mediafiles/documentation_records/Auto-Tune_Pro_Manual.pdf) (Zugriff: 16.01.2019).

Apple Inc. (2018): Logic Pro X, Pitch Correction – Übersicht, <https://help.apple.com/logicpro/mac/10.4/#/lgcef2835dcc> (Zugriff: 21.01.2019).

Celemony Software GmbH (2019): Melodyne 4, Handbuch. URL: <http://helpcenter.celemony.com/hc-2/pdf/melodyneStudio4/de?env=standAlone> (Zugriff: 20.01.2019).

## Artikel von Webseiten

Anderson, Lessley (2013): Seduced by 'perfect' pitch: How Auto-Tune conquered pop music. URL: <https://www.the-verge.com/2013/2/27/3964406/seduced-by-perfect-pitch-how-auto-tune-conquered-pop-music> (Zugriff: 13.02.2019).

BBC News (2010): X Factor admits tweaking vocals. URL: <https://www.bbc.com/news/entertainment-arts-11056050> (Zugriff: 09.02.2019).

Bernsee, Stephan (1999): Time Stretching and Pitch-Shifting of Audio Signals – An Overview. URL: <http://blogs.zynaptiq.com/bernsee/time-pitch-michael-r-h-overview/> (Zugriff: 13.01.2019).

Dolson, Mark (2014): The Phase Vocoder: A Tutorial. URL: <http://www.panix.com/%7Ejens/pvoc-dolson.par> (Zugriff: 13.01.19).

- Dolson, Mark; Laroche, Jean (1999): New phase-vocoder techniques for pitch-shifting, harmonizing and other exotic effects. URL: <http://www.ee.columbia.edu/~dpwe/papers/LaroD99-pvoc.pdf> (Zugriff: 15.01.2019).
- Marshall, Owen (2014): A Brief History of Auto-Tune. URL: <https://soundstudies-blog.com/2014/04/21/its-about-time-auto-tune/> (Zugriff: 08.02.2019).
- Savage, Mark (2010): How commonplace is auto-tune? URL: <https://www.bbc.co.uk/news/entertainment-arts-11056840> (Zugriff: 08.02.2019).
- Sengpiel Audio (1994): Formanten prägen die Klangfarbe. URL: <http://www.sengpielaudio.com/FormantenPraegenDieKlangfarbe.pdf> (Zugriff: 11.01.2019).
- The Daily Telegraph (2012): Sydney bound Justin Bieber admits voice tuned by software Melodyne for third album Believe. URL: <https://www.dailytelegraph.com.au/entertainment/sydney-bound-justin-bieber-admits-voice-tuned-by-software-melodyne-for-third-album-believe/news-story/43672f02f2103f812990e1232f1b36f5> (Zugriff: 09.02.2019).
- Thornton, Mike (2018): Antares Announces Auto-Tune Pro to Replace Auto-Tune 8. URL: <https://www.pro-tools-expert.com/home-page/2018/1/25/antares-announces-auto-tune-pro> (Zugriff: 18.01.2019).

## Videos

- Genius (2017): The History Of Auto-Tune | Genius News. URL: <https://www.youtube.com/watch?v=oS4Gku99Qco> (Zugriff: 09.02.2019)
- VOCALOID (2018): VOCALOID5 – Walkthrough. URL: <https://www.youtube.com/watch?v=UAtVGHI1AFM&t=27s> (Zugriff: 13.02.2019).
- Celemony (2011): Melodyne inventor Peter Neubäcker (film portrait). URL: <https://www.youtube.com/watch?v=u573PyXo-pY> (Zugriff: 20.01.2019).

## Herstellerseiten

- Ableton: URL: <https://www.ableton.com>



Antares Audio Technologie: URL: <https://www.antarestech.com>

Apple Logix Pro X: URL: <https://www.apple.com/de/logic-pro/>

Celemony Software GmbH: URL: <https://www.celemony.com/de/service1/about-celemony/company>

Yamaha Corporation, VOCALOID: URL: <http://www.vocaloid.com/en/>

Alle Internetquellen waren am 18.02.2019 erreichbar.

# Anhang

## Hörbeispiele

Beispiel 1: Ernst Toch – Fuge aus der Geographie (1930)

Beispiel 2: David Seville – The Chipmunk Song (1958)

Beispiel 3: Iannis Xenakis – Analogique A et B (1959)

Beispiel 4: T-Pain – Getcha Roll On

Beispiel 5: Yung Lean – Red Bottom Sky

Beispiel 6: Cher – Believe (Refrain)

Beispiel 7: Cher-Effekt – Auto-Tune-Simulation

Beispiel 8: Cher-Effekt – Melodyne-Simulation

Beispiel 9: Believe – ohne Auto-Tune Live

Beispiel 10: Believe – mit Auto-Tune Live (natürlich)

Beispiel 11: Believe – mit Auto Tune Live (Effekt)

Beispiel 12: Amazing Grace – ohne Korrektur

Beispiel 13: Amazing Grace – Melodyne-Korrektur

Beispiel 14: Amazing Grace – Auto-Tune-Korrektur

Beispiel 15: Amazing Grace - Stimmduplikation

Beispiel 16: Sailing – Tonhöhenverschiebung mit Ableton (-1)

Beispiel 17: Sailing – Tonhöhenverschiebung mit Auto-Tune (-1)

Beispiel 18: Sailing – Tonhöhenverschiebung mit Melodyne (-1)

Beispiel 19: Sailing – Tonhöhenverschiebung (-2)

Beispiel 20: Sailing – Tonhöhenverschiebung (+1)

Beispiel 21: Sailing – keine Tonhöhenverschiebung

**Anmerkung:** Alle Beispiele befinden sich in gleicher Weise aufgelistet auf dem mit eingereichtem USB-Stick.

## Umfrage

### Befragte Personen

Angela: 57 Jahre, Ärztin, Hobbymusikerin (singt in 3 Chören), hört privat selten Musik.

Anke: 46 Jahre, Sängerin, hört privat selten Musik.

Bianca: 32 Jahre, Erzieherin, hört privat gerne und oft Musik.

Elisa: 21 Jahre, Studentin (Medienmanagement), hört privat gerne und oft Musik.

Frank: 60 Jahre, Chorleiter, hört privat gerne und oft Musik.

Lynn: 24 Jahre, Studentin (Sonderpädagogik), hört privat selten Musik.

Martin: 48 Jahre, Sänger, hört privat selten Musik.

Paul: 23 Jahre, Student (Medientechnik), Hobbymusiker (Gitarrist einer Band), hört privat gerne und oft Musik.

Sven: 25 Jahre, Student (Psychologie), hört privat gerne und oft Musik.

Tom: 23 Jahre, Nachrichtenredakteur (Rundfunk), hört privat gerne und oft Musik.

## Ergebnis

Tonhöhenkorrektur: Welches Beispiel ist besser intoniert bzw. harmonischer?

Beispiel-Nr.	Programm	Anzahl Stimmen
13	Melodyne	5
14	Auto-Tune	4
-	Kein Unterschied	1

Tonhöhenverschiebung – Teil 1: Welches Beispiel klingt am natürlichsten?

Beispiel-Nr.	Programm	Anzahl Stimmen
16	Ableton	3
17	Auto-Tune	4
18	Melodyne	2
-	Kein Unterschied zw. Auto-Tune und Melodyne	1

Tonhöhenverschiebung – Teil 2: Bei welchem Beispiel handelt es sich um die nicht veränderte Originalaufnahme?

Beispiel-Nr.	Verschiebung in Halbtönen (mit Melodyne)	Anzahl Stimmen
18	-1	2
19	-2	1
20	+1	2
21	0	5

Cher-Effekt: Welches Beispiel kommt dem Effekt auf Chers Stimme näher?

Beispiel-Nr.	Programm	Anzahl Stimmen
7	Auto-Tune	6
8	Melodyne	5

## **Eigenständigkeitserklärung**

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

---

Ort, Datum

Vorname Nachname