

Musikalische Wahrnehmung und musikalisches Gedächtnis

Autor: Clemens Renner
Seminar: Musikalische Wahrnehmung
Dozent: Prof. Dr. G. Rötter
Sommersemester 2003

Inhalt

1	Schwerpunkt Wahrnehmung Echtheit der Wahrnehmung, Elementare Wahrnehmungs- ereignisse, Hörbereich und jnd, Die ADSR-Hüllkurve, Einfluss des Gedächtnis auf die Wahrnehmung	2
2	Schwerpunkt Gedächtnis Konstrukte zur Strukturierung, Sensorisches Gedächtnis	8
3	Beobachtungen Gestaltgesetze, Skalenillusion, Lokalisation der Schallquelle, Selektive Aufmerksamkeit, Mozart-Effekt	11
	Quellenverzeichnis, Abbildungsverzeichnis	15

1 Schwerpunkt Wahrnehmung

Echtheit der Wahrnehmung

Wer über die Wahrnehmung der Umwelt nachdenkt, wird sich schnell mit Philosophie konfrontiert sehen. Zum Beispiel drängt sich die Frage auf, ob das, was wir wahrnehmen, wirklich dem entspricht, was uns als Reiz von außen dargeboten wird. Oder vielleicht ist es wichtiger, darüber nachzudenken, wie stark wir das, was wir als Reiz empfangen, kategorisieren, favorisieren, verzerren oder direkt ablehnen und uns damit den Zugang zu einem Teil unserer Umwelt unbewusst versperren. Wie weit ist unsere Vorgeschichte für unsere Wahrnehmung verantwortlich? Welche Rolle spielt die Kultur, in der wir aufgewachsen sind?

Daneben könnten ganz andere Zusammenhänge von Bedeutung sein. So wäre es sicher interessant, eine eventuell vorhandene Beziehung zwischen der Qualität des musikalischen Gedächtnis und der Musikalität einer Person aufzuzeigen. Hierbei würde man für die Bestimmung der Musikalität der entsprechenden Person sicher einen Musikalitätstest heranziehen (vgl. [1]).

Betrachten wir zunächst ein einfaches Modell der Kommunikation:

Sender > Codierung der Mitteilung > Kanal > Decodierung der Mitteilung > Empfänger

Dieses Modell beschreibt den Weg der Nachrichtenübermittlung von einem Mitteilenden zu einem Empfangenden. Wer eine Nachricht übermitteln will, muss sich also zwangsläufig (nach diesem Modell) für eine spezifische Art der Übermittlung entscheiden. Er wählt einen Kanal zur Informationsübertragung aus. Als direkte Folge dieser Wahl muss sich der Sender auch schon für eine bestimmte Codierung der Information entscheiden. Wählt er beispielsweise als Medium zur Übermittlung ein Buch, so ist er damit automatisch Zwängen und Möglichkeiten ausgesetzt, die dieses spezifische Medium mit sich bringt. Er setzt so zum Beispiel beim Empfänger die Fähigkeit zu lesen voraus und muss sich damit abfinden, dass unter Umständen durch die fehlende Auszeichnung der Betonung ein Teil der intendierten Botschaft verloren geht oder zumindest verzerrt wird. Andererseits bietet ihm das Medium Buch auch die Möglichkeit, Abbildungen, Fußnoten, etc. einzubringen, um Sachverhalte klarer und vielleicht vollständiger zu erfassen. Denken wir nun einen Schritt weiter: Der Sender hat also eine mitzuteilende Botschaft und entscheidet sich für ein bestimmtes Medium der Informationsübermittlung. Geht man davon aus, dass jedes Medium Nachteile mit sich bringt, was zumindest für die meisten gegeben sein wird, so ist die Konsequenz aus dem eben Geschilderten, dass praktisch keine Nachrichtenübermittlung verzerrungsfrei und unverfälscht erfolgen kann. Dies hat natürlich eine direkte Auswirkung auf unsere Wahrnehmung, heißt das doch, dass wir nie genau das wahrnehmen, was uns der Sender mitteilen will. Da diese Überlegungen aber eher philosophischer Natur sind, werden wir uns im Rahmen dieser Ausarbeitung nicht weiter damit beschäftigen. Es sei aber angemerkt, dass dieses Phänomen durchaus beachtenswert ist und zumindest berücksichtigt werden sollte, wenn man über die Wahrnehmung der Umwelt spricht.

Im Folgenden werden wir uns einiger Wahrnehmungsprozesse widmen, wie sie Bob Snyder formuliert hat. Da ich Snyders Theorien für mich selbst erst übersetzen musste, um sie zu verstehen, erfolgt natürlich auch hier wieder vielerlei Verzerrung und Einbringung von Unschärfe, einmal durch mein Verstehen von Snyders Thesen und dann in der Übersetzung und gleichzeitiger Neuformulierung im Deutschen.

Elementare Wahrnehmungsereignisse

Zunächst hat Snyder einige elementare Wahrnehmungen beschrieben, um daraus Konstrukte entwickeln zu können, die die musikalische Wahrnehmung erfassen können. So spricht er von *Ereignisverschmelzung* oder *Ereignisfusion* (*event fusion*), wenn unsere Wahrnehmung Ketten von eintreffenden Schwingungen bezogen auf Frequenz oder Lautstärke kategorisiert und somit zusammenfasst. Das Ergebnis dieser Wahrnehmung findet sich kurzzeitig im echoischen Gedächtnis wieder.

Das Kernstück seiner Theorie aber ist ein „Ereignis mit Tonhöhe“, von ihm *pitched event*, von mir im folgenden *Tonhöhenereignis* genannt. Darunter versteht Snyder die kleinste Organisationseinheit in der Wahrnehmung von Melodien. Ein Tonhöhenereignis besteht aus einer Menge von ähnlichen, akustischen Schwingungen, die schnell (genug) aufeinander folgen, dass kein Rhythmus erkannt und das Wahrgenommene zu einer einzigen Tonhöhe zusammengefasst wird. Allerdings ist eine direkte Abgrenzung zwischen Tonhöhenereignis und Geräuschereignis nur schwer möglich. Im Allgemeinen befindet sich der Klang eines Instruments dabei auf einer gleitenden Skala zwischen Tonhöhe (etwa Streicher und Holzbläser) und Geräusch (Schlagzeug oder perkussive Instrumente). Nun stellt sich die Frage, wann eine Folge von Ereignissen schnell genug ist. Snyder spricht hier von etwa fünf bis zwanzig Ereignissen pro Sekunde. Die Anzahl der ankommenden akustischen Schwingungen wird in einem gewissen Zeitabschnitt immer höher, was dazu führt, dass sie nicht mehr einzeln unterscheidbar sind und anders verarbeitet werden. Warren stellte in einer Untersuchung 1999 fest, dass diese Schwelle bei etwa fünf Ereignissen pro Sekunde liegt, man also bis zu dieser Grenze noch zwei nicht-identische, aufeinander folgende Ereignisse unterscheiden kann. Wird die Grenze von etwa zehn Ereignissen pro Sekunde überschritten, greift die so genannte holistische Wahrnehmung, also die ganzheitliche Betrachtung von Information. Sie sorgt dafür, dass wir eher einen Klang, ein Geräusch, also mehr den Charakter einer Wellenform hören als den eines Tons (Warren, 1993). Spätestens bei einer Informationsdichte von 20 Ereignissen pro Sekunde jedoch werden alle Ereignisse, die praktische Verwendung finden, zu einer Wellenform zusammengefasst und somit zu einem Ereignis auf einer höheren Ebene der Wahrnehmung. Ohr und Gehirn werden also mit zunehmender Dichte der Information mehr und mehr Klang- als Ereignisverarbeiter. Snyder illustriert diesen Zusammenhang in Abbildung 1. In der linken Spalte finden sich Frequenzen im Sinne von Ereignissen (oder Zyklen, Perioden bei Schwingungen) pro Sekunde. Daneben steht dementsprechend als Reziproke die Dauer eines Ereignis in Sekunden. Die Schwelle der Ereignisfusion ist im unteren Bereich kenntlich gemacht, sie wird bei 8 bis 16 Ereignissen pro Sekunde angesetzt. Auf der rechten Seite wiederum ist eine Entsprechung der Ereignisfrequenzen in Tonhöhen angegeben. So würde eine Frequenz von 512 Hz etwa dem Ton h entsprechen, also leicht oberhalb des Kammertons a liegen (440 Hz).

Da wir nun wissen, wie wir eine Tonhöhe zu verstehen haben, ist die nächste interessante Größe das *Intervall*, das Snyder als den melodischen Abstand zwischen zwei Tonhöhen fasst. Soweit folgt dies auch der gängigen Definition des Begriffs. Eine Tonhöhe wird höher, eine tiefer wahrgenommen, den Abstand bezeichnen wir als Intervall. Eine an dieser Stelle vielleicht noch nicht unmittelbar verständliche Unterscheidung von großen und kleinen Intervallen soll sich später noch als nützlich herausstellen. Nehmen wir vorerst kleine Intervalle als *Stufen* (*steps*) und große als *Sprünge* (*leaps*) an. Die nächsthöhere Organisationseinheit ist die *Melodie*. Snyder sieht sie als Folge unterschiedlicher Tonhöhenereignisse. Auch diese Definition folgt dem intuitiven Verständnis des Begriffs.

Kommen wir noch einmal auf die Schwelle der Ereignisfusion zurück. Während wir unterhalb dieser Schwelle Rhythmen direkt wahrnehmen und verarbeiten können, ist uns dies mit zunehmender Anzahl von Ereignissen pro Sekunde nicht mehr möglich. Offensichtlich gibt es aber

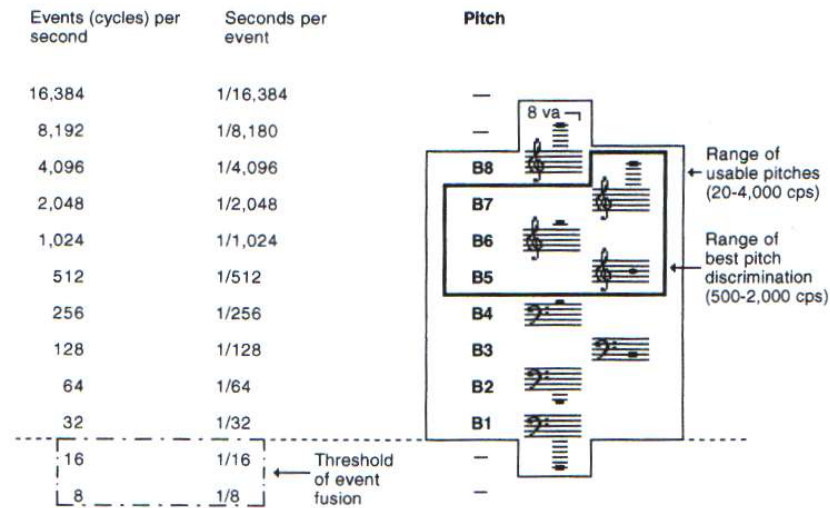


Abb. 1: Zeit und Tonhöhe, Ereignisfusion

auch oberhalb dieser Schwelle rhythmische Phänomene, die von Snyder als Mikrorhythmen bezeichnet werden, da sie zu fein sind, um direkt wahrgenommen werden zu können. Sie gehen jedoch nicht unter: Was wir im Langsamen als Rhythmus hören, manifestiert sich bei höherer Ereignisdichte im Timbre bzw. in der Tonfärbung. Die rhythmischen Muster von Ereignissen finden sich also in einem Mikrorhythmus wieder, der zu schnell dargeboten wird, um ihn wie langsamere Rhythmen verarbeiten zu können.

Hörbereich und jnd

Das menschliche Gehör nimmt bekanntlich Frequenzen zwischen etwa 20 Hz und 20 kHz wahr. Der für die Unterscheidung von Tonhöhen interessanteste Bereich liegt dabei zwischen 500 und 2000 Hz. In diesem Zusammenhang spricht man von *just noticeable distance* oder auch *just noticeable difference (jnd)*, also dem kleinsten bemerkbaren (oder besser: wahrnehmbaren) Unterschied zwischen zwei Tonhöhen. 1982 fanden Barlow und Mellon heraus, dass der Mensch etwa 0,3-prozentige Abweichungen von zwei Frequenzen feststellen kann. Zwischen 500 und 2000 Hz finden sich zwei Oktaven wieder, denn eine Verdopplung der Frequenz entspricht der Verschiebung eines Tons um eine Oktave nach oben. Analog kann man mit Frequenzhalbierung eine Oktave nach unten gehen.

Zwischen etwa 20 Hz und 4 kHz liegt der Bereich der praktisch nutzbaren Tonhöhen, in denen die klare Unterscheidung zweier verschiedener Frequenzen stattfindet. Außerhalb dieses Bereichs (etwa ab vier bis fünf kHz) werden die Töne dann eher heller als höher empfunden, je weiter man Richtung obere Hörgrenze läuft, während in der anderen Richtung Töne mit Frequenzen unter 20 Hz vornehmlich einem Brummen gleichen, zunehmend „körniger“ klingen und sich letztendlich in einzelne Impulse aufsplitten. Auch für den an absoluten Zahlen interessierten Leser sei hier eine genannt: Unter experimentellen Bedingungen stellte Handel 1989 fest, dass mindestens 1400 verschiedene Frequenzen vom menschlichen Gehör unterscheidbar sind.

Die ADSR-Hüllkurve

Wie wir vorhin gesehen haben, werden Schwingungen, die schnell genug aufeinander folgen, dabei

eine konstante Rate aufweisen und eine ähnliche Wellenform haben, als eine Tonhöhe wahrgenommen. Im Gegensatz dazu stehen verhältnismäßig unstrukturierte Klänge ohne interne, periodische Struktur. Diese werden als „weißes Rauschen“ bezeichnet und ohne Tonhöhe wahrgenommen. Erhöht man hier die Frequenz, wird das Rauschen als heller empfunden, mit sinkender Frequenz klingt es dumpfer. Betrachtet man den Klang eines Instruments, so findet sich dort eine Mischung aus Geräusch und Ton, da nahezu alle Instrumentenklänge eine so genannte ADSR-Hüllkurve aufweisen. Man betrachtet dabei den Schwingungsverlauf in vier Phasen: *Attack* (Anschlag), *Decay* (Dämpfung), *Sustain* (Haltepegel) und *Release* (verklingen, ausklingen). Jedes Instrument hat eine charakteristische Hüllkurve. So ist beim Klavier ein recht markanter Anschlag zu finden, während es praktisch keine Sustain-Phase gibt, da der Ton nach dem Anschlagen sofort beginnt zu verklingen, Decay und Release gehen damit ziemlich direkt ineinander über. Im Gegensatz dazu findet sich beim Klang einer Orgel nahezu kein Dämpfungs- und Ausklinganteil. Der Ton schwingt ein, bleibt relativ konstant auf dem Haltepegel und verstummt auch sofort, wenn die Taste losgelassen wird, der Schall, der dann noch zu hören ist, stammt ja meist vom Hall der die Orgel umgebenden Kirche. Abb. 2 illustriert die Form einer Hüllkurve (*envelope*).

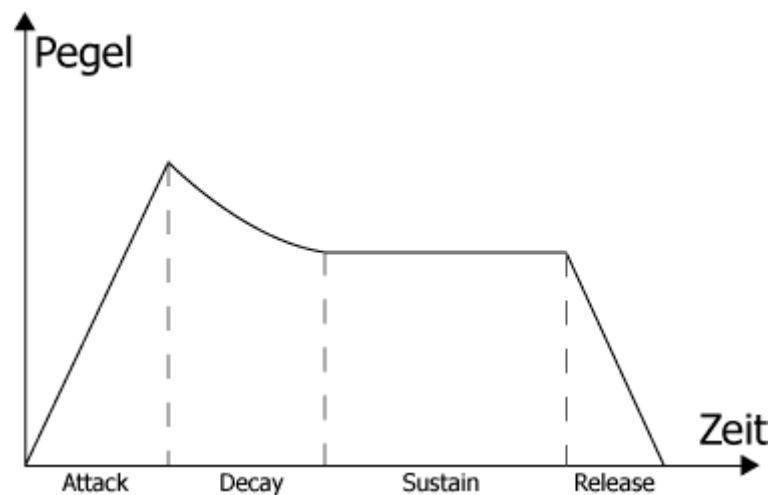


Abb. 2: Die Phasen einer ADSR-Hüllkurve

Wir laufen nach rechts auf der Zeitachse, beginnend bei Null. Attack- und Decay-Phase zusammengenommen (Anschlag und Anschlagsdämpfung) bilden den Einschwingvorgang. Danach kommt der Sustainpegel zur Geltung, mit dieser Intensität nämlich wird der Ton gehalten, während die Release-Phase das Ausschwingen verkörpert. Wie bereits erwähnt, kann die Sustainphase zum Beispiel nur aus einem direkten Übergang von Decay zu Release bestehen (Klavier), da der Ton nach dem Einschwingen direkt wieder verklingt.

Einfluss des Gedächtnis auf die Wahrnehmung

An dieser Stelle möchte ich zum ersten Mal auf einen wesentlichen Anteil des Gedächtnis an der Wahrnehmung zu sprechen kommen. Solange der Vergleich von Tonhöhen über das echoische Gedächtnis statt findet, ist die Unterscheidung einfach, da die Wahrnehmung hier noch am genauesten ist und der geringsten Abstraktion unterliegt. Sobald jedoch das Kurzzeitgedächtnis bemüht werden muss, weil der zeitliche Abstand zwischen zwei Tönen zu groß wird, werden eher konzeptuelle Kategorien verglichen als reine Tonhöhen. Das Problem ist recht simpel: Um mit der Masse an aufzunehmenden und zu behaltenden Informationen fertig zu werden, können wir nicht alle Details erinnern und müssen uns demnach auf Ausschnitte des Wahrgenommenen beschränken. Zu diesem Punkt führte Deutsch 1999 eine Untersuchung durch, in der gezeigt werden konnte, dass zwei zeitlich direkt benachbarte Töne (also ein Intervall) einfacher unterschieden werden können, als ein durch andere, effektiv störende Töne unterbrochenes Intervall; genauer: man nimmt ein Intervall und unterbricht es durch eine Folge von Tönen. Hier wäre ein Ansatzpunkt, die Gestaltgesetze auf die musikalische Wahrnehmung zu beziehen, da offensichtlich die zeitliche Nähe eine Rolle spielt. Im Zuge der Abstraktion, die durch den Vergleich im Kurzzeitgedächtnis entsteht, werden eventuell Unterschiede fortgelassen, die bei direkt aufeinander folgender Darbietung aufgefallen wären. Anschaulich lässt sich das vielleicht so formulieren: Wenn ein Chor singt, lässt es sich oft nicht vermeiden, dass man leicht aus der Stimmung gerät und absolut gesehen, häufig etwa einen Halbton tiefer endet, als man begonnen hat. Ist beispielsweise erster und letzter Ton des Stücks ein A, so ist es oft am Ende doch ein As, was dem Zuhörer aber für gewöhnlich nicht auffällt, da der Ton im Vergleich zu den anderen Stimmen in der Harmonie liegt und in der Regel nicht direkt gegen das „richtigere“ A erklingt. Hätte der Chor nacheinander A und As gesungen, wäre der Unterschied sicher aufgefallen.

Die Wahrnehmung ist also wesentlich durch das Kurzzeitgedächtnis beschränkt: man behält nur wenige einzelne Elemente und merkt sich eher Schemen. Ebenso werden Vergleiche eher auf Schemen und Kategorien durchgeführt und weniger auf den direkten Repräsentationen des Wahrgenommenen.

Bevor wir uns jedoch weiter mit dem Gedächtnis und dessen Einfluss auf die Wahrnehmung beschäftigen, kommen wir noch einmal auf die Intervalle zurück. Es wurde bereits erwähnt, dass Snyder eine Unterscheidung zwischen Stufen (steps, kleiner Frequenzabstand) und Sprüngen (leaps, großer Frequenzabstand) vornimmt. Hören wir benachbarte Stufen (zum Beispiel bei einer Tonleiter), so ist uns auch gegenwärtig, dass es sich um aneinander liegende Stufen handelt. Die Abstände zwischen den Tonhöhenereignissen sind jedoch auch so groß, dass die Unterscheidung der einzelnen Töne problemlos möglich ist. Unsere Wahrnehmung geht dabei in vielen Fällen davon aus, dass Töne, die von der selben Schallquelle stammen, auch in einem gewissen Rahmen bleiben und keine großen Sprünge machen. Anschaulich wird dies klar, wenn man über die menschliche Sprechstimme nachdenkt. Sie bewegt sich meist nach oben und unten wenig abweichend um eine zentrale Frequenz, eben die „Höhe“ der Sprechstimme. Während wir dazu neigen, benachbarte Intervalle zu einer Kontur zusammenzufassen und aus ihnen eine Gruppe zu bilden, dienen gerade die Sprünge zur Abgrenzung dieser Konturen, sie fassen diese Konturen sozusagen rahmend ein. Bemüht man die Gestaltgesetze, speziell das Gesetz der Nähe, so wird deutlich, dass eine engere Bindung zwischen benachbarten Tönen (also bei Stufen-Intervallen) besteht als zwischen den Tönen eines Sprungs.

In diesem Zusammenhang ist auch die Transposition von Melodien beachtenswert. Obwohl die absoluten Tonhöhen sich ja offensichtlich unterscheiden, wenn man eine Melodie beispielsweise nach unten oktaviert, so ist doch die innere Struktur, also das, was die Melodie eigentlich ausmacht,

immer noch erkennbar. Solange also die Intervalle stimmen, ist die absolute Tonhöhe nahezu vernachlässigbar.

Und da die Oktavierung gerade zur Sprache kam: obwohl die Oktave an sich nach Snyders Definition ein Sprung ist, sind sich zwei Tonhöhen, die eine Oktave voneinander entfernt sind, doch so ähnlich, dass man sie als nahezu gleich empfindet. Eine mögliche Erklärung könnte man mit der Struktur der Obertöne ansetzen. Jeder natürliche, also nicht elektronisch generierte Klang, beispielsweise der eines Streichinstruments setzt sich aus vielen Frequenzen zusammen. Die Tonhöhe, die man hört, entspricht zwar der Frequenz des Grundtons dieses Klangs, es schwingen jedoch in einem von Instrument zu Instrument verschiedenen Spektrum andere Frequenzen mit. So hat der Klang einer Geige ein anderes Obertonspektrum als der einer Flöte oder der eines Klaviers. Der erste Oberton ist dabei eben genau der Grundton mit doppelter Frequenz, also der um eine Oktave nach oben transponierte Grundton. Dieser Erklärungsansatz bringt aber schon bei wenig näherer Betrachtung ein Problem mit sich. Die Frequenzen der Obertöne steigen nur logarithmisch an, das heißt, der zweite Oberton liegt eine Quinte über dem ersten, der dritte Oberton ist eine Terz höher als der zweite, usw. Warum soll nun der erste Oberton bedeutsamer sein als die anderen?

Es gibt ein sehr viel ausgereifteres und anschaulicheres Modell zur Erklärung dieser vom Frequenzabstand eigentlich nicht verständlichen Tonverwandtschaft. Das Chroma (wie in Abb. 3 veranschaulicht) teilt das hörbare Frequenzspektrum in zwölf verschiedene Ton- oder Frequenz-

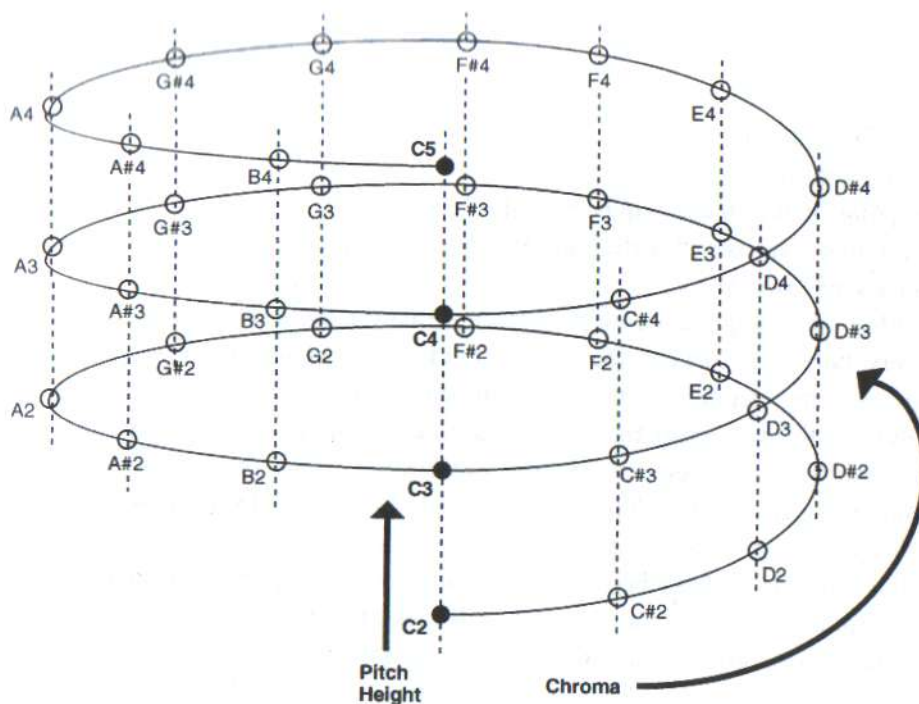


Abb. 3: Visualisierung des Chroma

kategorien ein. Diese lassen sich problemlos am Klavier beobachten, gibt es doch den mit C bezeichneten Ton sieben mal, genauso wie den Ton A. Töne sind sich also ähnlich, entweder wenn sie in der Frequenz ähnlich sind oder in ihrem Namen. Es gibt also frequenzverschiedene Töne, die sich im Chroma ähneln, ebenso wie man chromaverschiedene Töne findet, die in der Frequenz nahezu gleich sind. Das Chroma erklärt damit in Form einer Spirale die Verschiedenartigkeit der steigenden Frequenzen, da sie auf der Wanderung entlang des Spiralbogens immer weiter von einer tieferen Frequenz entfernt sind. Andererseits finden sich auf einer vertikalen Linie immer die

oktavgleichen Töne, so dass sich einmal die Nähe in Form von kleinen Intervallen in dem Modell wiederfindet (die Spirale weiter hinauf oder hinab wandern), genauso aber die Verwandtschaft von Tönen, die zwölf Halbtöne auseinander liegen (eine Vertikale hinauf oder hinab wandern). Die Entwickler dieses Modells, Dowling und Harwood (1986) beobachteten auch, dass die meisten Stimmsysteme auf Oktaven basieren, nicht nur unser westliches. Begründet haben sie diesen Effekt damit, dass die verschiedenen Frequenzen als gleich empfunden werden, wohl unabhängig von der kulturellen Herkunft der Hörenden. Die Oktave scheint auch noch eine weitere Geschlossenheit aufzuweisen, da uns Intervalle, die eine größere Spanne als die der Oktave abdecken, tendenziell eher als Sprünge erscheinen (Deutsch, 1999). Wir haben also Mühe, sie in einer Gruppe zu sehen, da Sprünge ja nach Snyders Definition als Grenze einer Gruppe fungieren.

Im Folgenden werden wir uns einige weitere Konstrukte Snyders ansehen, die eher das Gedächtnis beschreiben als die Wahrnehmung.

2 Schwerpunkt Gedächtnis

Konstrukte zur Strukturierung

Unstrittig ist sicher, dass man einen zeitlichen Verlauf oder eine über die Zeit veränderliche Größe nur echt wahrnehmen kann, wenn man die Änderung beobachten kann. Dies setzt jedoch offensichtlich voraus, dass man zumindest zwei verschiedene Ereignisse miteinander vergleichen kann, die nicht gleichzeitig aufgetreten sind. Die logische Konsequenz daraus ist die Notwendigkeit eines Gedächtnis, um eine frühere Information mit einer neueren zu vergleichen. Sehen wir uns Snyders Definition einer *Melodie* an: er versteht darunter eine Folge hörbarer Ereignisse, die erkennbare Muster einer Kontur aufweist. Zwischen den Einzelereignissen müssen Intervalle wahrnehmbar sein. Für den Hörer ist es an dieser Stelle vor allem wichtig, dass er eine Einteilung in Höhe und Tiefe vornehmen kann. Daraus wiederum lässt sich schließen, dass nicht nur eine Folge von Tönen eines Instruments eine Melodie darstellen kann, sondern diese auch einfach aus unterschiedlich „hohem“ Rauschen bestehen kann.

Als wir uns am Ende des letzten Abschnitts mit den Eigenarten der Oktave beschäftigt haben, fiel der Begriff *Stimmsystem*, auf den ich nun zurück komme. Snyder versteht darunter Folgendes: Ein abstraktes, erzeugendes System von Intervallen, das grundlegende Intervallkategorien festlegt, so insbesondere auch ein kleinstes Intervall. Als gemeinsamer Nenner verschiedener Stimmsysteme findet sich wieder das Intervall der Oktave, da die meisten Stimmsysteme Intervalle verwenden, die kleiner als eine Oktave sind. Was wir als Tonleiter kennen, findet sich in Snyders Begriffswelt als *Skala* wieder. Eine Skala bedient sich dabei einer Teilmenge der insgesamt zur Verfügung stehenden Frequenzen und Intervalle eines Stimmsystems. Möglicherweise gibt es in einer Skala auch eine zentrale Frequenz oder Tonhöhe. Während die Anzahl der Elemente eines Stimmsystems quasi beliebig ist, zeigt das Kurzzeitgedächtnis die Grenzen bei der Menge der Teile einer Skala auf. So ist es also wenig erstaunlich, dass wir in unseren Tonleitern acht Töne wiederfinden, ist doch die Anzahl von „Brocken“ (*chunks*), die unser Kurzzeitgedächtnis aufnehmen kann durch fünf bis neun Elemente begrenzt. Man könnte noch weitergehen und den Umfang der Tonleiter auf sieben Töne einschränken – wobei wir genau bei den magischen 7 ± 2 Chunks landen – schließlich sind der erste und der letzte Ton gleich. Daneben ist es durchaus möglich, dass Skalen mit Emotionen identifiziert werden; Wir kennen dieses Phänomen wohl am ehesten von der Unterscheidung zwischen Dur (leicht und optimistisch) und Moll (eher melancholisch oder schwer).

Abbildung 4 verdeutlicht den Unterschied zwischen den Intervallen eines Stimmsystems und denen einer Skala. Während die elementaren Intervalle im Stimmsystem alle gleich groß sind (linke Seite), finden sich in der Skala (oder Tonleiter) Intervalle verschiedener Größe, die eben gerade diese bestimmte Skala kennzeichnen.

Sind diese Stimmsysteme nun natürlich oder künstlich? Sie sind erlernt. Wenn jemand ohne musikalische Vorbildung ist, also keine Harmonielehre und Gehörbildung bekommen hat, wird er bei einer ihm hörbar gemachten Terz wohl nicht sagen, dass es sich dabei um eine Terz handelt. Er wird vielmehr ein Intervall hören, also zwei Töne, die in einem gewissen Abstand voneinander liegen. Betreibt man eine musikalische Ausbildung mit ihm, wird er sicher auch ein relatives Gehör entwickeln und die Intervalle benennen können. Konfrontiert man hingegen jemanden, der ein bestimmtes Stimmsystem verinnerlicht hat, mit einem anderen, so wird ihm diese andere Einteilung in Kategorien nicht als solche erscheinen, sondern das neue System wird zunächst „verstimmt“ wirken, im Gegensatz zu dem, das er schon kennt. Nicht zuletzt findet sich das Stimmsystem auch in

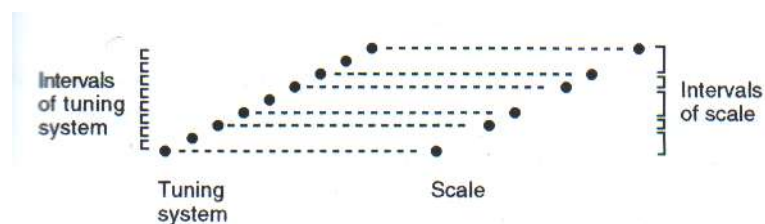


Abb. 4: Intervalle in einem Stimmsystem und in einer Skala

der Beschaffenheit der zu diesem System gehörenden Instrumente wieder. So ist es mit einem Klavier doch nur möglich, „schiefe“ Töne zu spielen, wenn das Klavier selbst verstimmt ist. Es ist also, wenn Instrument und Stimmsystem zueinander passen, tendenziell einfacher, in der Stimmung zu bleiben und schwerer, davon abzuweichen.

Snyder stellt nun die folgende Frage: Wenn unser Hören in Kategorien funktioniert, warum hören wir dann schiefe Töne? Hören wir einen Klang, der von mindestens zwei Instrumenten ausgeht, so ist es einfach, eine sehr kleine Abweichung festzustellen, wenn beide Instrumente den selben Ton spielen. Es kommt dabei zum Phänomen der Schwebung. Es tritt auf, wenn sich zwei Frequenzen überlagern, die in ihrer Periodenlänge nur marginal verschieden sind. Obwohl natürlich nicht jeder diese Erscheinung benennen kann, so erkennt sie doch jeder sofort. Man empfindet die Töne als verschieden. Scheinbar ist es also leichter, Töne verschiedener Klangquellen miteinander zu vergleichen und für „gerade“ oder „schief“ zu befinden, als dies für eine einzige Klangquelle zu tun. Letztendlich, das behaupte ich, ist auch das Hören eines Tons unter Umständen ein Hören verschiedener Klangquellen, zumindest bei erfahrenen Musikern, da diese einen Vergleichston in sich tragen. Sie haben eine Vorstellung davon, wie der Ton klingen soll und vergleichen Sollwert mit Istwert.

Wir ergänzen nun die Definition des Stimmsystems um folgenden Punkt: Ein Stimmsystem definiert lediglich die Intervalle, es nimmt keine hierarchische Einteilung dieser vor. Im Wesentlichen haben wir damit also einen Baukasten, mit dem man die Musik eines bestimmten Kulturkreises erzeugen kann, nämlich dessen, zu dem das Stimmsystem gehört, bzw. aus welchem es stammt.

Um der vielleicht noch etwas einseitigen Definition etwas mehr Leben einzuhauchen, erwähnt Snyder die indonesische Gamelan-Musik. In ihr gibt es auch ein Stimmsystem, doch dies ist unserem recht unähnlich, denn es erlaubt Dialekte. Nur innerhalb eines Orchesters sind die Intervalle des Stimmsystems gleich, ein anderes Orchester variiert die Intervalle und formt damit eine abweichende Färbung des Systems. So ist es theoretisch wohl möglich, ein Konzert mit diesen beiden Orchestern zu veranstalten, praktisch werden die eigenen Ohren aber wahrscheinlich lieber nur einem Orchester lauschen, um in dessen Dialekt zu bleiben.

Um sich nun an ein Stück Musik erinnern zu können, gibt es mehrere Möglichkeiten: Man hört es oft oder man besorgt sich eine Aufnahme des Stücks, die einem das Erinnern nun im Wesentlichen abnimmt, kann man sich ihrer doch bedienen, um eine ziemlich akkurate Reproduktion des Stücks zu erleben. Eine dritte Möglichkeit besteht in der schriftlichen Verfestigung der Musik. Man kann ein Stück Musik ja dadurch konservieren, dass man es in Noten festhält. Liest man es wieder, wird man unter Umständen an eine echte Darbietung erinnert. Ist nun aber eine Wiedergabe einer Sinfonie, die in Form von Notenmaterial vorliegt, durch ein Orchester *echt*? Oder sind es gerade die Abweichungen von der Notation, die das Stück in Bezug auf das Wiedererkennen ausmachen? Man denke an dieser Stelle vornehmlich an bewusste Verzerrungen. So wird jeder Berufsmusiker, der ein Streichinstrument spielt, das Vibrato beherrschen. Im Sinne der Echtheit der Wiedergabe des Notenmaterials ist dies natürlich eine Verfälschung, sie ist jedoch so gering, dass sie den eigentlichen Ton nicht so stark verzerrt, als dass er nicht mehr erkannt würde. In der Regel spricht man hier von bis zu 0,5-prozentigen Abweichungen (z.B. in der Frequenz), die jedoch die Wahrnehmung nicht stören. Ein anderes Beispiel, besonders aus dem Jazz bekannt, sind die *blue notes*, also das bewusste Verfälschen des Ansatzes eines Tons, indem man diesen tiefer ansetzt, als er notiert ist, ihn jedoch dann „heraufschleift“, letztendlich die eigentliche Tonhöhe noch erreicht. So ist es doch vielleicht gerade das Ausmaß oder die Art der Verfälschung, was wir uns merken und woran wir uns erinnern. Im Allgemeinen jedoch geht man davon aus, dass Nuancen eher schlecht behalten werden, da die feinste Wahrnehmung und Erinnerung, wie bereits erwähnt, im echoischen Gedächtnis stattfindet. Das vergrößerte und verallgemeinerte Merken greift demnach auch hier. So kann ein wiederholtes Anhören eines Stücks immer wieder Neues mit sich bringen, weil es nicht in allen Details gespeichert wird. Mir ist jedoch auch das genau gegenteilige Phänomen bekannt. So passiert es durchaus, dass ich ein Stück gerade wegen verschiedenster Nuancen immer wieder höre, mich jedes mal darauf gefasst mache, genau diese Feinheiten wieder zu hören und mich daran zu erfreuen.

Nachdem wir nun vornehmlich Snyders Sicht der Dinge und sein Erklärungsmodell kennen gelernt haben, werden im letzten Teil vornehmlich losgelöst davon einige andere Ansätze und Beobachtungen vorgestellt.

Das sensorische Gedächtnis

Was geschieht ganz am Anfang der musikalischen Wahrnehmung? Nach de la Motte-Haber (1989) setzt hier zuerst das sensorische Gedächtnis an. Es stellt hier die Art der ankommenden Information fest und trifft gegebenenfalls bereits erste Entscheidungen über dessen Bedeutung. Die Zuordnung von gehörter Musik zu einer bestimmten Stilrichtung basiert danach auf der Aktivität des sensorischen Gedächtnis. In diesem Zusammenhang ist interessant, dass sogar trotz Maskierung eine Wahrnehmung eines Reizes mit entsprechender Reaktion erfolgen kann. Unter Maskierung wird hier verstanden, dass der eigentliche Reiz durch einen anderen Reiz (z.B. durch ein Geräusch) überdeckt wird, der ursprüngliche Reiz also eigentlich nicht mehr wirklich ankommt. Dies bedeutet aber, dass wir auch ohne bewusste Wahrnehmung in der Lage sind, darauf zu reagieren. Dies ist schon bemerkenswert, impliziert es doch, dass wir zumindest ein Stück weit nicht Herr unserer selbst sind.

3 Beobachtungen

Gestaltgesetze, Skalenillusion, Lokalisation der Schallquelle

Die Übertragung der Gestaltgesetze auf die Musik wurde schon einmal angesprochen, sie soll an dieser Stelle etwas weiter ausgeführt werden. Dazu stellen wir einmal die holistische (ganzheitliche) und die analytische Wahrnehmung gegenüber. Das Gesetz der Nähe findet seine Entsprechung im *auditory streaming*, der Gruppierung frequenznaher Töne zu einem melodischen Muster. Auch in Snyders Begriffswelt ist dies wiederzufinden, wenn man seine Definition von Melodie ansieht (Folge unterschiedlicher Tonhöhenereignisse, jedoch nur in Intervallen von Stufen, Sprünge bilden eher den Rahmen um eine Melodie). Auf der anderen Seite steht die *stream segregation*. Dieser Effekt tritt in der Regel auf, wenn zwei hoch-tief-alternierende Folgen von Tönen dargeboten werden. Ist das Tempo ausreichend, wird unter Umständen eine zumindest pro Stimme nicht vorhandene Melodieführung wahrgenommen (siehe Abb. 5).

Es werden also die Töne, die in ihrer Frequenz nahe beieinander liegen, zu Melodien gruppiert. Einerseits bietet sich hier das Gesetz der Geschlossenheit zur Erklärung an, wonach (auch Snyders

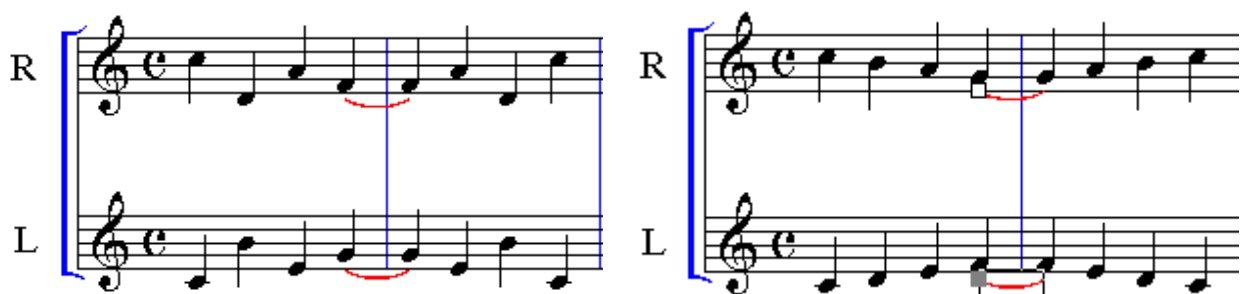


Abb. 5: Eigentlich dargebotene Tonfolgen (links) und wahrgenommene Tonfolgen (rechts)

Definition folgend) die großen Intervalle eher als Grenze fungieren und daher die kleineren Intervalle den Vorzug bekommen und eine Melodie formen. Andererseits gibt es noch das Gestaltgesetz des gemeinsamen Schicksals, das hier scheinbar greift: die beiden losen Tonfolgen werden als zwei Linien mit gleicher Bestimmung gehört.

Reicht das Tempo der Darbietung hingegen nicht aus, so bleibt mehr Zeit für die detailgenaue Analyse des Gehörten, der Schwerpunkt liegt dann auf der Wahrnehmung der Teile und man verfällt nicht so leicht der Skalenillusion. Von Davidson, Power und Michie gab es dazu 1987 eine Untersuchung, in der sie die Illusion an drei Gruppen testeten. Die Komponisten (erste Gruppe) hörten keine Skalen, sondern konnten die einzelnen Ströme (*streams*) isolieren, während die Musiker (zweite Gruppe) der Illusion zum Teil verfielen. Ein Ansatz zur Begründung war hier, dass das Arbeitsgedächtnis anscheinend auf vier Töne begrenzt ist, denn die ersten vier Töne konnten sie noch richtig auseinander halten, danach griff der Effekt der gehörten Tonleiter. Die letzte Gruppe der musikalischen Laien brachte das für mich erstaunlichste Ergebnis: spielte man ihnen vorher den linken und den rechten Kanal getrennt vor, konnten sie richtig diskriminieren. Die Begründung: Sie hatten eine Erwartung entwickelt, was nun kommen würde – und das, was sie im Anschluss hörten, passte in dieses Schema. Daher fiel ihnen die Aufteilung leichter.

Wir hören also sowohl analytisch als auch ganzheitlich. Die Wahrnehmung ist weder sequentiell, sodass das eine auf das andere folgt, noch findet sie parallel statt. Es liegt nahe, hier von zwei unterschiedlich funktionierenden Arten der Wahrnehmung zu sprechen (de la Motte, [3]).

Eine wirklich interessante Beobachtung machte Diana Deutsch 1994. Sie stellte fest, dass beim Musikhören ein elementarer Prozess der Wahrnehmung außer Kraft gesetzt wird. Wir können über minimale Laufzeitunterschiede feststellen, aus welcher Richtung uns ein Schall erreicht. Braucht er etwas länger zum linken Ohr als zum rechten, wissen wir, dass die Schallquelle rechts von uns liegt, da der Schall dort früher ankam und nicht erst um den Kopf herum fließen musste (was er tut, um das linke Ohr zu erreichen). Beim Musikhören, fand Deutsch, wird das Richtungshören außer Kraft gesetzt. Sie begründete diesen Effekt damit, dass das Hören von Musik ein räumlicher Prozess ist, bei dem vornehmlich Tiefe und Breite eine Rolle spielen. Das räumliche Vorstellungsvermögen ist daher so vereinnahmt, dass es dadurch unfähig wird, den Ursprung des Schalls festzustellen. In der Tat ist diese Eigenart gehörter Musik essentieller Bestandteil in der Tontechnik, speziell bei Abmischungen im Studio. Es gibt dort durchaus die Möglichkeit, eine Mischung von Aufnahmen zu erstellen, bei der einige Stimmen weiter im Vordergrund erscheinen (z.B. ein Soloinstrument), während andere eher im Hintergrund bleiben (begleitendes Orchester). Auch die Breite des entstehenden Klangs kann beeinflusst werden, je nachdem, wie man die einzelnen Tonspuren auf den linken oder den rechten Stereokanal verteilt.

Selektive Aufmerksamkeit

Wie wir bereits gesehen haben, ist das Kurzzeitgedächtnis nur mit begrenzter Kapazität ausgestattet. Ankommende Reize werden zusammengefasst, beispielsweise zu chunks oder Gestalten. Allerdings wird diese Gruppierung zu bedeutsamen Einheiten nur mit einem Teil der ankommenden Informationen vorgenommen. Ein Erklärungsmodell besagt, dass hierbei eine Hemmung zu beobachten ist, der unwesentliche Teil wird also ausgefiltert. Alternativ spricht man von einem begrenzten Fassungsvermögen, was einfach durch seinen beschränkten Platz dafür sorgt, dass nur Bestimmtes behalten wird, während irrelevante Informationen verloren gehen. Man stelle sich die zur Verfügung stehende Aufmerksamkeit wie einen Topf vor, aus dem alle anstehenden Aufgaben einen Teil abbekommen. Es gibt nur endlich viele Ressourcen, so dass alle Aufgaben (oder *tasks*) damit auskommen müssen. Die Aufmerksamkeit kann sich also nur aufteilen. Das klingt zunächst nicht sonderlich effizient, wird aber dadurch entschärft, dass wir mit erlernten Fertigkeiten dem Gedächtnis eine Menge Arbeit abnehmen können und somit einen Teil der Aufmerksamkeit ausgliedern und damit weniger Kapazität für diese Task hergeben müssen. Sehen wir uns aber mit

mehreren, ähnlichen Aufgaben konfrontiert, die alle gleichsam anspruchsvoll sind, so wirkt dies kontraproduktiv: die ähnlichen Reize engen die zur Verfügung stehende Konzentration ein (Pechmann, Mohr, 1992). So können wir problemlos musikalische und visuelle Reize miteinander vereinbaren, zum Beispiel, wenn wir Musik hören und dabei ein Bild betrachten (musikalisch/visuell), oder Musik hören und dabei einem Sprecher zuhören, der vielleicht etwas über den Komponisten des Stücks oder das Stück selbst erzählt (musikalisch/sprachlich). Hingegen bereiten uns zwei gleichzeitig gehörte Musikstücke Kopfschmerzen. Bei dem Versuch, sich auf eines von beiden zu konzentrieren, schwindet die Aufmerksamkeit für das jeweils andere, beide gleichsam konzentriert zu verfolgen, scheint unmöglich (musikalisch/musikalisch).

Die Methode des Superlearning versuchte sich diesen Ansatz zu nutze zu machen. Man versprach sich durch paralleles Musikhören während einer anderen (primären) Tätigkeit eine angehobene Aufmerksamkeit und folglich mehr mögliche Konzentration für die primäre Aktivität. Bringt man das in Einklang mit der beobachtbaren, begrenzten Kapazität des Kurzzeitgedächtnis, so ist die Folge, dass es sich bei der Musik nur um solche handeln darf, die in ihrer Struktur einfach ist. Vielleicht ist das auch der Grund, warum die Musik, die in den Charts läuft und sich so gut verkauft, simple genug ist, um sie nebenbei hören zu können.

Ebenso gut, wie man die begrenzte Kapazität als Ursache für eine Selektion ansehen kann, lässt sich auch die Umkehrung konstruieren. Man nehme an, dass eine Selektion stattfindet, so ist die Konsequenz, dass nur eine begrenzte Kapazität zur Verfügung steht, da nicht alles, was ankommt, durchgelassen wird.

Der Mozart-Effekt

Als Letztes möchte ich noch auf ein Forschungsergebnis (vgl. dazu [4], [5]) zu sprechen kommen, dass 1993 für Aufsehen in der Fachwelt gesorgt hat. Forscher aus den USA hatten ein Experiment durchgeführt, dessen Ergebnis lautete: Hören von Mozartmusik die Intelligenz steigert. Warum gerade Mozart? Präzisieren wir zunächst etwas: zum einen zeigten die Versuchspersonen keine höhere Intelligenz, sondern „lediglich“ ein wesentlich besseres Abschneiden in einem Intelligenztest, bei dem bestimmte, kognitive Fähigkeiten gemessen wurden. Von 36 College-Studenten hörte ein Drittel den Anfang von Mozarts Sonate für zwei Klaviere (KV 448, Gruppe A), das nächste Dutzend genoss Entspannungsmusik vom Band (Gruppe B) und die letzten zwölf Studenten befanden sich zehn Minuten lang in Stille (Gruppe C). Direkt im Anschluss an diese „Aufwärmphase“ unterschiedlichen Charakters folgte ein Teil des Stanford-Binet-Intelligenztests, bei dem es in einem Teil um das räumliche Vorstellungsvermögen ging. Die Probanden sollten dabei in der Vorstellung einen Scherenschnitt anfertigen, d.h. ein Blatt Papier gedanklich mehrfach falten, mit einer Schere bearbeiten und die beim Aufklappen entstehende Form beschreiben. Der Test dauerte etwa 10-15 Minuten und hier zeigte sich die Disposition der Mozarthörer. Die Probanden aus Gruppe A zeigten im räumlichen Test signifikant höhere Leistungen als die Vergleichspersonen aus den anderen Gruppen.

Die Untersuchung wurde von Gordon Shaw (Physiker) und Frances Rauscher (Spezialist auf dem Gebiet der kognitiven Entwicklung) durchgeführt und zumindest Shaw verdient heute Geld mit Produkten, die sich auf die nicht reproduzierbaren Resultate dieser Untersuchung stützen.

Es drängen sich folgende Fragen auf: Zeigt sich der Effekt nur bei diesem Mozartstück oder auch bei anderen Werken des selben Komponisten? Vielleicht sogar bei anderen Komponisten? Oder

liegt es generell nur daran, dass vorher Musik gehört wurde? Tatsächlich kann man diese Fragen nur so beantworten: Bereits geringfügige Veränderungen des Versuchs machten den Effekt zunichte und ließen die Gruppe A nicht mehr herausragen. Die wahrscheinlich wichtigste Frage ist doch aber: lassen sich diese Ergebnisse reproduzieren? Letztere kann mit einem klaren Nein beantwortet werden.

Motiviert, diesem Phänomen auf den Grund zu gehen, gab es einige glaubwürdigere Studien, die sich mit dessen Ergründung beschäftigten. Man geht davon aus, dass die Musik eine positive Wirkung hat. Nun ist dies nichts Neues, haben wir doch eine ähnliche Argumentation schon beim Superlearning gesehen. Allerdings wird hier auf die „Parallelbeschallung“ verzichtet und lediglich das höhere psychische Aktivierungsniveau ausgenutzt. Dies erklärt jedoch noch nicht, warum der Effekt nicht bei anderer Musik auftritt. Mit Hilfe von Computern wurden Mozart-Partituren durchsucht, um festzustellen, was ihnen gemein ist. Das Ergebnis ist wenig überraschend: es zeigt Merkmale der Frühklassik auf, eben gerade der Periode, der Mozarts Musik zuzuordnen ist. Genauer:

- **Natürlichkeit:** Es handelt sich um Dur, die Melodik ist oft aus Dreiklängen abgeleitet. Die melodischen Wendungen haben Fanfarencharakter, man findet eine einfache Harmonik. Es wurde Quintverwandschaft beobachtet.
- **Symmetrie:** Die musikalische Form gründet sich auf geschlossenen Themen (Perioden) und in kleinen wie großen Formteilen herrscht Ausgewogenheit und Gleichheit.

Man deutet dies wie folgt: Die Fanfaren haben Signalwirkung, man fühlt sich angesprochen und wird aufmerksam. Die Konturen, die präsentiert werden, sind einfach nachzuvollziehen und animieren zum Mitsingen oder -summen. Die Periodik bringt nun noch Ordnung in das klangliche Geschehen, die Gruppierung und Strukturierung fällt leicht. Letztendlich gleicht die wiederkehrende Gegenüberstellung von Dominante und Tonika einem Frage- und Antwortspiel, das immer gleich ausgeht. Man wird dazu verleitet, Vorhersagen zu treffen, die dann auch erfüllt werden. Der erfahrene Erfolg verlangt nach Fortsetzung – man geht dementsprechend motivierter an den nachfolgenden Test.

Die Ergebnisse des Experiments wurden anfangs nur in einigen Absätzen in einem wissenschaftlichen Journal erwähnt, von der Fachwelt sehr kritisch aufgenommen und mittels mangelnder Reproduzierbarkeit als nicht nachvollziehbarer Unfug abgestempelt. Nach John Bruer gibt es für den beobachteten, nicht nachweisbaren Effekt „keine soliden Fundamente in der Neurowissenschaft“ ([5]). So bleibt es dem gemeinen Musikwahrnehmer nur, seiner Wahrnehmung nicht bedingungslos zu trauen und auch nach kritischen Stimmen zu suchen. Denn, wie eingangs erwähnt, allein die Wahl des zur Kommunikation benutzten Mediums hat unter Umständen schon Einfluss auf den übermittelten Inhalt.

Quellenverzeichnis

- [1] Renner, Clemens: Was ist Musikalität?, Dortmund, 2003.
[2] Snyder, Bob: Music and Memory: an introduction. Cambridge, 2000.
[3] laMotte-Haber, Helga de: Handbuch der Musikpsychologie. Laaber, 1996.
[4] <http://www.mensch-und-musik.at/forschung/mozarteffekt>
Anmerkung: Unglücklicherweise kann diese Adresse nicht mehr abgerufen werden, ich stieß jedoch beim Versuch, eine alternative Adresse zu finden, auf eine sehr viel kritischere und allgemein mir objektiver erscheinende Betrachtung des Zusammenhangs (Quelle [5]). Zudem ist sie sehr unterhaltsam zu lesen.
[5] <http://www.skeptischeecke.de/Worterbuch/Mondphasen/Mozart-Effekt/mozart-effekt.html>

Abbildungsverzeichnis

- Abb. 1: [2], S. 125, Fig. 10.1.
Abb. 2: http://www.netzmarkt.de/thomann/bilder/synthesizer_adsr_kurve.gif
Abb. 3: [2], S. 131, Fig. 10.2.
Abb. 4: [2], S. 139, Fig. 11.1.
Abb. 5: <http://www.borg-krems.ac.at/~kboucher/akust.illusionen/skalenillusion.gif>
Abb. 6: <http://www.borg-krems.ac.at/~kboucher/akust.illusionen/skalenwiewahrgenommen.gif>

Der Autor ist per E-Mail zu erreichen: email@clemens-renner.de .