

# BILDERWELT DER KLÄNGE — KLANGWELT DER BILDER. BEOBACHTUNGEN ZUR KONVERGENZ DER SINNE

Helmut Rösing

Schenkt man den Äußerungen so mancher Kulturkritiker Glauben, dann leben wir in einem »optischen Zeitalter«: In den letzten Dezennien sei eine Generation von Augenmenschen herangewachsen, für die der Umgang mit den audiovisuellen Medien zu den Selbstverständlichkeiten des Alltags gehört (vgl. Pech 1969, Hausheer/Schönholzer 1994). Bei der Wahrnehmung würden generell die optischen gegenüber den nonverbal-akustischen Informationen dominieren. Damit bewahrheite sich, was Rudolf Arnheim bereits in seiner 1932 erschienenen Schrift *Film als Kunst* argwöhnte: Filmmusik sei vor allem dann gut, wenn man sie nicht bemerke (zit. n. Arnheim 1974: 304), und was Siegfried Kracauer in seiner *Theorie des Films – Die Errettung der äußeren Wirklichkeit* (1964: 210) meinte, als er schrieb: »Das Medium Film, so scheint es, kann der Musik keine Hauptrolle gestatten und muß daher automatisch ihren Prioritätsanspruch zurückweisen.«

Weiter noch geht der Vorwurf, Musik könne von der Mehrzahl der Konsumenten heutzutage ohnehin kaum unabhängig von den visuellen Botschaften rezipiert werden, mit denen sie im Film oder Videoclip verknüpft seien. Die optische Brille habe sogar dort Bestand, wo die visuelle Botschaft – etwa auf dem CD-Filmmusik-Sampler – ausbleibt; visuelle Zwangskonditionierung sei der Normalfall. Frank Zappa brachte das 1992 in einem ARD-Interview in Frankfurt am Main auf die griffige Formel: Für die Mehrheit der Bevölkerung sei Musik mittlerweile vor allem das, was über die Bildschirme flimmere.

Wie bei allen derartigen Verallgemeinerungen ist Skepsis angebracht. Legt man den Musikkonsum der auditiven Medien – Tonkassette, CD, musikbezogene Internetnutzung – als Indikator für die aktuellen Wahrnehmungsgepflogenheiten zugrunde, so ließe sich mit dem gleichen Recht auch von einem akustischen Zeitalter sprechen: Das signalisiert die intensive Nutzung des Hörfunks als einem auditiven Medium par excellence z.B. ebenso wie die allerorten praktizierte Beschallung der Alltagswelt mit Lautsprecher-

musik – ganz unabhängig von der Frage, ob das nun der Wertigkeit von Musik eher zu- oder abträglich sei.

Doch allem Anschein nach ist es gar nicht sinnvoll, derart streng zwischen auditiver und visueller Wahrnehmung zu trennen und die eine Wahrnehmungsmodalität gegenüber der anderen derart auszuspielen, wie das z.B. Norbert Jürgen Schneider (1997) in seinem Filmhandbuch tut, indem er auf den »polaren Gegensatz« von Auge und Ohr abhebt. Schließlich haben wissenschaftliche Untersuchungen seit den Anfängen der Synästhesieforschung im ausgehenden 19. Jahrhundert immer wieder deutlich werden lassen, dass menschliche Wahrnehmung generell auf der Komplementarität von Auge und Ohr beruht, d.h. intermodal angelegt ist (im Überblick: Rösing 1998a). Bilderwelten der Klänge und Klangwelten der Bilder sind das Ergebnis. Das sei im Folgenden etwas näher beleuchtet. Neurophysiologische und rezeptionspsychologische Hinweise zur Konvergenz der Sinne sollen die Grundlage für eine Theorie der audiovisuellen Musikwahrnehmung geben und erste Anhaltspunkte dafür, wie eine Filmmusik- oder Videoclip-Analyse beschaffen sein könnte, die nicht nur objektimmanent, sondern auch wahrnehmungsorientiert ist.

## **1. Neurophysiologische und rezeptionspsychologische Hinweise**

Bei Beispielen für intermodale Wahrnehmung wird wiederholt auf archaische Kulturen mit ihren ganzheitlichen Lebensformen und die direkte Verknüpfung von Musik und Kult verwiesen, aber ebenso auf das alte China (heteromodale Zuordnung von Tönen und Farben zu Jahreszeiten, Himmelsrichtungen, Elementen und Musikinstrumenten) und auf die griechische Antike (Prinzip der kosmischen Weltordnung und der Kongruenz von Zahlenverhältnissen, Intervallen, Zusammenklängen mit der Harmonie der Sphären). Aristoteles postulierte bereits um 350 v.u.Z. in Zusammenhang mit seiner Theorie der Assoziationen das Vorhandensein eines übergeordneten Sinns, der die Wahrnehmungen der einzelnen Sinne koordiniere (vgl. Hurte 1982).

Die universalästhetisch-philosophischen Konzeptionen der Antike waren auch für das europäische Mittelalter und die beginnende Neuzeit verbindlich. Einen Höhepunkt fanden sie in den Schriften des Jesuiten Athanasius Kircher (1650). Er versuchte die Identität von Licht und Schall auf der Grundlage von übereinstimmenden Zahlenverhältnissen bzw. Intervallen zu erklären. Ein erster Ansatz, diese Zusammenhänge naturwissenschaftlich-

physikalisch zu untermauern, stammt von Isaac Newton. In seinen *Opticks* von 1704 konstruierte er eine Analogie zwischen den sieben Spektralfarben und den sieben Tönen der diatonischen Skala. Diese mechanistische Zuordnung wurde von Johann Wolfgang Goethe in seiner Farbenlehre (1791) strikt abgelehnt, während sie Louis-Bertrand Castel (1725) zur Grundlage der Konstruktion eines Farbenklaviers machte (vgl. Jewanski 1999).

In Verbindung mit der Synästhesieforschung wurde die Frage nach der Intermodalität der Wahrnehmung auch für die Musikwissenschaft aktuell. In den 1920er Jahren erschienen Arbeiten, die die spezielle Synästhesieproblematik, also die zwangsweise Verbindung von Tönen mit Farben bzw. Formen beim musikalischen Rezeptionsprozess in einen größeren anthropologisch-kulturgeschichtlichen Zusammenhang stellten. Gerade die Beschäftigung mit der Musik schriftloser Ethnien führte Erich Moritz v. Hornbostel (1924: 519) sehr bald zu der Einsicht: »Es gibt übersinnliche Wahrnehmungen. Bewegung kann man sehen, hören oder tasten«. Und hinsichtlich der Qualität der Helligkeit konstatierte er, sie sei »nicht nur eine analoge, sondern eine identische Seite der Phänomene der verschiedenen Sinnesgebiete«. Dementsprechend propagierte nur wenig später Albert Wellek mit seinen Untersuchungen zum Doppelempfinden in der abendländischen Kulturgeschichte (seit 1928) eine Theorie der »Urentsprechungen«. Der Psychologe Heinz Werner gab dafür konkrete Beispiele (1966): Bewegung, Raum, Zeit, Intensität, Rauigkeit sind ihm zufolge intersensorielle Qualitäten. Ihre umfassende Wahrnehmung sei auf eine Konvergenz der Sinne angewiesen, unbeschadet der Tatsache, dass jeder Sinn ein anderes Segment von Umwelt wahrnehme und somit über einen jeweils spezifischen Leistungsvorsprung verfüge.

Die wichtigsten Ergebnisse der fachwissenschaftlichen Literatur zur intermodalen Wahrnehmung hat der Psychologe und Kulturwissenschaftler Lawrence E. Marks (1978: 5ff.) in fünf Punkten zusammengefasst.

- (1) Die einzelnen Sinne sind aus einem Ursinn hervorgegangen und haben sich im Verlauf der Menschheitsgeschichte zwar immer weiter ausdifferenziert, nicht aber ihre gemeinsame Basis verloren;
- (2) die verschiedenen Sinne können über ein- und dieselbe Gegebenheit der Umwelt informieren;
- (3) den Wahrnehmungsobjekten selbst haften intersensorielle Qualitäten an, die bei ihrer Rezeption zu subjektiven wie auch gesellschaftlich geprägten Analogiebildungen und Assoziationen führen;

- (4) die verschiedenen Sinne verfügen über vergleichbare Prinzipien der Informationsverarbeitung auf der Grundlage psychophysikalischer Gesetzmäßigkeiten; und
- (5) es bestehen neuronale Korrespondenzen bei der internen Verarbeitung von Informationen aller Sinneskanäle im Gehirn.

Insbesondere den Punkten 4 und 5 galt die neurophysiologische Forschung ab den 1980er Jahren. Das Gehirn nimmt die Umwelt wahr, interpretiert sie und steuert das Verhalten des Organismus. Die Nervenzellen des Gehirns reagieren ausschließlich auf elektrochemische Signale. Jede über die Sinnesrezeptoren dem Gehirn zur weiteren Verarbeitung zugeführte Information muss folglich in elektrochemische Impulse umgewandelt werden. Einer Zellerregung lässt sich nicht entnehmen, ob sie von einer Licht- oder Schallwelle, einem Geruchsmolekül oder einer mechanischen Einwirkung herrührt. Es gilt das »Prinzip der Neutralität neuronaler Codes« (Roth 1997: 80). Entscheidend ist, was das Gehirn als selbstreferentielles System aus diesen unspezifischen Codes macht.

Für das Entstehen von Wahrnehmungsinhalten sind zumindest fünf Eigenschaften der Reizstrukturen ausschlaggebend (vgl. Roth 1997: 108): die Modalität – z.B. akustisch oder optisch; die Qualität innerhalb einer Modalität – Tonhöhe und Klangfarbe; die Intensität – Reizstärke; die Zeitstruktur – Anfang, Dauer und Ende des Reizes; der Ort – Lokalisation im Raum. Deren Verarbeitung im Gehirn erfolgt auf dreierlei Weise: konvergent, parallel und divergent. Konvergenz betrifft die Zusammenfassung und Mischung von Information verschiedener Modalität, Parallelverarbeitung meint die Verarbeitung bestimmter Qualitäten einer Information auf getrennten Nervenbahnen und in unterschiedlichen Hirnarealen, Divergenz beinhaltet die Aufspaltung der ursprünglichen zur Schaffung von neuer Information. Das führt zu einer immer größeren neuronalen Aktivität in Bezug zum Ausgangsreiz (Roth 1997: 121ff.). So konstruiert das Gehirn aus eher spärlich zu nennenden peripheren Reizkonstellationen bzw. Eingangsdaten der Sinnesorgane die Wahrnehmung. Das jedoch geschieht keineswegs ausschließlich nach intersubjektiv identischen Wahrnehmungsmustern. Denn die frühere Annahme einer Verarbeitungshierarchie von den niederen zu den höheren Verarbeitungszentren hat sich als falsch erwiesen. Generell leiten nicht nur afferente Nervenbahnen im bottom-up-Verfahren Signale an höhere Hirnebenen weiter, sondern auf umgekehrten Weg – top-down – transportieren afferente Bahnen Befehle zu den unteren Hirnarealen. Auf diese Weise können auf Vorerfahrung und individueller Sozialisation beruhende Gedächtnisinhalte, also Wissen und Gefühle (emotionale Bewertung) die Verarbeitung

von Information bereits auf der untersten neuronalen Ebene beeinflussen – bis hin zu dem Extremfall, dass Wahrnehmung blockiert wird, weil sie unbekannt ist und/oder als bedeutungslos eingestuft wird.

Auch von der Annahme, dass die interne Verarbeitung zum überwiegenden Teil über die sensorspezifischen Nervenbahnen und in speziellen Hör- und Sehzentren des Gehirns erfolge, gilt es sich zu verabschieden. Die Schaltstellen bzw. Nuclei des Nervus acusticus z.B. ermöglichen eine effiziente Parallelverarbeitung. So führen vom Colliculus inferior des Zwischenhirns Nervenverbindungen zum Kleinhirn (Auslösung von Reflexen) und zum Sehnerv. Das erlaubt die Abgleichung mit Informationen des visuellen Systems. Auch im Nucleus posterior des Thalamus treffen Signale von Auge und Ohr zusammen und können konvergieren. Der Thalamus wiederum projiziert seine Erregungen nicht nur in die spezifischen sensorischen Rindengebiete, sondern darüber hinaus in die kortikalen Assoziationsareale. Hier werden u.a. intermodale Verknüpfungen unter Einfluss der Bewertungs- und Gedächtnissysteme des limbischen Systems weiterverarbeitet. Die hier skizzierte funktionelle Zusammenschaltung verschiedener, gerade auch nichtspezifischer Cortexareale beim Hören von Musik ist durch EEG-Mapping-technik mittlerweile bestens belegt (Altenmüller 2000). Somit folgt: Der Wahrnehmungsprozess mit seinen verschiedenen Stufen der internen Verarbeitung bis hin zur kognitiv realisierten Wahrnehmung beruht auf weit mehr als allein auf der Grundlage der über einen Sinnesrezeptor eingehenden Reize. Er ist das Ergebnis des Miteinanders der eingegangenen Reizstrukturen mit all dem, was im Langzeitspeicher des Gehirns (Gedächtnis) bereits als Erfahrungsinventar (Wissen) existiert. Wahrnehmung ist darum nur in besonderen Fällen ausschließlich akustisch oder optisch ausgerichtet. Üblicher Weise kommt es zur Konvergenz der Sinne und zu einer intermodalen Zusammenschau. Das führt nahezu zwangsläufig zur Theorie der audiovisuellen Musikwahrnehmung.

## **2. Theorie der audiovisuellen Musikwahrnehmung**

Jede Art des Musikhörens, die über die rein vegetative Rezeption hinausgeht, ist ein aktiv-schöpferischer Vorgang. Alle in einer aktuellen Situation wahrgenommenen musikalischen Schallereignisse stehen in direkter Wechselbeziehung mit dem bisherigen Musikkonzept und mit dem allgemeinen Erfahrungsinventar einer Person, mit emotionalen, assoziativen, kognitiven Schemata und Prototypen. Die jeweilige Sozialisation (mit personen- und gesellschaftsabhängigen Variablen wie Alter, Geschlecht, Familie, Ausbil-

dung, soziales Milieu, Medien) ist von entscheidender Bedeutung dafür, wie ein erklingendes Stück Musik bewertet bzw. verstanden und in den eigenen Erfahrungshorizont integriert wird (Rösing 1997). Somit liegt die Annahme nahe, dass eine strikte Trennung in auditive und visuelle Wahrnehmung ein der Wahrnehmungsrealität nur bedingt gerecht werdendes Konstrukt darstellt. Musikalisches Erleben kommt meistens durch eine Koppelung von auditiven Reizen mit optischen Eindrücken bzw. Imaginationen zustande: Das führt zur Bilderwelt der Klänge und, darüber hinaus, zur Klangwelt der Sinne. Dabei liegt eine Unterscheidung in sechs verschiedenartige, empirisch durch Beobachtung und Befragung ermittelte audiovisuelle Wahrnehmungsmodi bzw. -ebenen nahe (s. bereits Rösing 1998b).

### **(1) Real musikbezogene Wahrnehmungsebene**

Vor dem Zeitalter der Übertragungsmusik war das Hören von Musik so gut wie ausschließlich an die Live-Darbietung gebunden. Musik aus Automaten zählte zu den ebenso bestaunten wie wohlwollend belächelten Ausnahmen. Musikerinnen und Musiker haben ihre Auftritte immer schon auf optisch attraktive Weise in Szene gesetzt, etwa durch die Anordnung auf dem Podium, durch typische Kleidung und Aufmachung, durch aufeinander abgestimmte Bewegungen beim Singen und Spielen, durch vielfache Show-Elemente (Salmen 1988). Nur in seltenen Fällen waren die Ausführenden den Blicken der Zuhörer entzogen. Aber zu sehen gab es selbst dann noch etwas: räumliches Interieur, andere Zuhörer, Tanzende.

### **(2) Imaginierte visuelle Wahrnehmungsebene**

Mit Beginn der Tonträger-Ära wurde das reale audiovisuelle Ganzheitserlebnis auf ein rein akustisches Teilerlebnis reduziert. Den Befürwortern der absoluten Musik kam diese Entwicklung durchaus entgegen. Als Pendant zum Notentext galt ihnen die von allen visuellen Störelementen bereinigte Tonträger-Fassung (Adorno 1976). Was die Musikpuristen aber übersahen, ist der Umstand, dass beim Wahrnehmen von Musik ohne die direkt dazugehörige optische Komponente der Musikedarbietung zumindest zweierlei geschehen kann. Entweder ergibt sich eine optische Verbindung mit einer nicht musikbezogenen, aber real gegebenen optischen Situation – dazu mehr unter Punkt 5 – oder aber es treten beim Hören, z.B. mit geschlossenen Augen, Bildvorstellungen auf. Schon Kurt Hubers grundlegende Untersuchungen aus der Anfangszeit der Musikpsychologie (1923) ergaben, dass derartige intermodale Wahrnehmungen bis hin zum »Sphärenenerlebnis« durchaus üblich sind. Vorstellungen bzw. Imaginationen haben meist einen Doppelcharakter. Sie enthalten willkürliche, kog-

nitiv steuerbare und unwillkürliche, sich der kognitiven Kontrolle entziehende Elemente. Auch sind sie nicht immer verbalisierbar. Darauf hat bereits Tibor Kneif (1970: 15ff.) in den 1970er Jahren hingewiesen. Eine Bestätigung geben die empirischen Untersuchungen von Günter Kleinen (1994) zur »psychologischen Wirklichkeit von Musik«.

### **(3) Auskomponierte Ebene visueller Vorstellungen**

Durch Musik hervorgerufene bildhafte Vorstellungen können sehr frei und vielfältig sein. Denn neben dem musikalischen Reiz als auslösender Größe sind viele intrasubjektive, z.B. durch kulturelle Traditionen oder auch den Zeitgeist beeinflusste Faktoren von Bedeutung. Dass es aber auch direkt musikbezogene, intersubjektive, im Kern invariante Vorstellungsinhalte gibt, zeigen die rezeptionsästhetischen Ausführungen von Hans Heinrich Eggebrecht (1994) zur Wirkungsgeschichte der Musik Beethovens ebenso wie z.B. die faktorenanalytische Auswertung von Hörprotokollen zur Musik Debussys, wie sie Michel Imberty (1971) vorgelegt hat. Demnach stehen den Komponisten bzw. Musikern vielfältige Möglichkeiten zur Verfügung, um durch das musikalische Material selbst eine inhaltliche Lenkung und Konkretisierung von Vorstellungsbildern beim Rezipienten zu bewirken. Ein Überblick über Tonmalerei und Programmmusik kann davon ein beeindruckendes Zeugnis ablegen (Klauwell 1968). Der von den Vertretern der Autonomieästhetik im 19. Jahrhundert erhobene Vorwurf, hier pfusche Musik mit dem Pinsel statt sich auf ihre eigene auditive Ebene zu beschränken, ist nicht erst seit Videoclip und Multimedialität fragwürdig. Schon Joseph Haydn hat sich bekanntlich in seinen Oratorien tonmalerische Möglichkeiten auf höchst publikumswirksame Art zunutze gemacht, um das Verbot szenischer Aufführungen (Oper, Ballett) zur Fastenzeit zu umgehen. Er komponierte die visuelle Ebene in die Musik. Was real auf der Bühne nicht gezeigt werden durfte, konnte sich so als visuelle Vorstellung in den Köpfen der Zuhörer entwickeln.

### **(4) In Szene gesetzte (inszenierte) visuelle Ebene**

Der Schritt von der auskomponierten zur tatsächlich inszenierten visuellen Ebene ist naheliegend und derart selbstverständlich, dass er dort, wo es um eine Systematik der Funktionen von Musik im Film oder Videoclip geht, kaum Erwähnung findet. Denn schon jede Bühnenmusik – Ballett, Schauspielmusik, Oper –, ebenso jede Tanz- oder Ritualmusik der schriftlosen Kulturen, bildet eine Einheit mit der szenischen Inszenierung. Und das gilt nicht nur dort, wo Aufführungsabläufe traditionell oder normativ festgelegt sind, sondern auch dort, wo mit jeder Inszenie-

rung wieder neue Bildebenen zur Musik ausgelotet werden wie bei der Oper, bei Performances in der Disco oder im Videoclip. Die Faszination einer in möglichst vielen Wahrnehmungsbereichen auskomponierten multimedialen Verknüpfung von intersensoriellen Vorstellungswelten kommt auf besonders intensive Weise im Gesamtkunstwerk zum Ausdruck. Es hat, von Wagners Neukonzeption des Musikdramas aus, einen Siegeszug ohnegleichen quer durch alle Stilrichtungen und Künste bis zu Happening, Performance, Rockspektakel angetreten (Rösing 1993).

#### **(5) Reale, nicht musikbezogene visuelle Wahrnehmungsebene**

Bei seiner Aufstellung einer Rangliste der Künste war für Kant (1799, § 53) der Umstand bedeutsam, dass man zwar die Augen schließen und wegsehen könne, nicht aber die Ohren, um wegzuhören. In der Realität allerdings dürfte das Wegsehen immer dann eher die Ausnahme sein, wenn es wirklich etwas zu sehen gibt. Das gilt auch im Fall des Musikhörens. Konzentriertes Hinhören mit geschlossenen Augen beschreibt lediglich eine vom Musikrezipienten zum Verhaltensritual der Innerlichkeit hochstilisierte Ausnahmesituation, gleichgültig, ob es sich um Live- oder um Übertragungsmusik handelt. Gerade die Verknüpfung von Musik mit musikfremden visuellen Eindrücken gehört zu einer besonders interessanten Konstellation. Schon lange vor dem technischen Zeitalter gab es Musik als Reisebegleitung – in der Postkutsche, auf dem Mississippi-Dampfer, beim Marsch. Mittlerweile befindet sich in nahezu jedem Auto ein Radio, und jeder ICE- oder Flugreisende kann sich Kopfhörer ausleihen. Ein Spezialfall musikalischer Rezeption ist zur Selbstverständlichkeit geworden, weil er den Strategien menschlicher Wahrnehmung offenbar sehr entgegenkommt. Die Welt, die man sieht und durch die man sich bewegt, rückt auf Distanz: Musik aus Kopfhörern und beim Reisen haucht der weitgehend um den Originalton gebrachten Außenwelt neues Leben ein. Daraus resultiert der Eindruck, die visuelle Seite der Umwelt laufe wie im Kino oder beim Ansehen eines Videoclips an einem vorüber (Schönhammer 1988). Dieses Phänomen der Spaltung von Sinnesräumen ist alt. Es war auch schon ein wesentliches Anliegen z.B. bei der Ausübung von Besessenheitsriten (Rouget 1985), bevor es die Neue Welt auf Disco-Partys, Tribal-Dances, Techno-Raves usw. eingeholt hat.

#### **(6) Filmische Visualisierungsebene**

Die Annahme, in Film und Videoclip werde der Musik nun endlich jene reale optische Ebene wiedergegeben, die sie spätestens als Lautsprechermusik eingebüßt hat, trägt. Die Performance-Ebene der Musikdarbie-

tung lässt sich allenfalls rudimentär in der filmischen Totale einfangen. Wie wenig zufriedenstellend das aber ist, haben die Diskussionen um das Genre Musikfilm zur Genüge deutlich gemacht. Das Abbild von Aufführungsrealität in Film oder Videoclip kann ebenso wenig mit dem Gesamtambiente einer Aufführung identisch sein wie eine Studioproduktion. Das hat Kurt Blaukopf (1989) mit dem Begriff der »Mediamorphose« anschaulich beschrieben. Weniger problematisch ist dagegen die filmische Realisation einer vorgestellten, imaginierten Ebene zur Musik. Was diesbezüglich in Filmen wie *Fantasia* von Walt Disney (USA 1940) oder *Yellow Submarine* von den Beatles (Großbritannien 1967) gemacht wurde, zählt heutzutage zum Produktionsstandard.

Eine wichtige Voraussetzung für die Akzeptanz filmisch realisierter interner Vorstellungswelten besteht darin, dass sich der Rezipient auf das jeweilige Bilderangebot einlässt, welches sich mit seinen eigenen Bildvorstellungen zur Musik ja keineswegs zu decken braucht. Dabei pflegt die qualitative Kluft zwischen filmischer Ins-Bild-Setzung und intrasubjektiven Vorstellungsbildern bei Musik mit deutlich auskomponierter visueller Ebene als besonders krass empfunden zu werden. Um es an einem klassischen Beispiel zu verdeutlichen: Eine optische Gewitterdarstellung zum musikalischen Gewitter in Beethovens »Pastorale« wäre plattes Mickey-Mousing. Ein realitätsbezogenes Abbild dürfte mit dem Reichtum der beim Hörer durch die Musik aktivierten Vorstellungsbilder in der Regel ebenso wenig mithalten können wie mit dem strukturellen und informativen Reichtum der Musik.

Widersetzt sich das Medium Musik der filmischen Vereinnahmung auf den Wahrnehmungsebenen 1 bis 4 in hohem Maß, so lässt sich aber ein deutlicher Bezug zur Musikrezeption auf Reisen oder mit dem Walkman (reale, nicht musikbezogene visuelle Wahrnehmungsebene) erkennen. In beiden Fällen handelt es sich, wie Klaus-Ernst Behne (1987: 183) zu Recht konstatiert, um eine hochgradig »artifizielle Wahrnehmungssituation«. Mit einem entscheidenden Unterschied: In Film und Videoclip wird die reale, nicht primär auf die musikalische Struktur bezogene visuelle Ebene durch eine spezielle, von den Filme- bzw. Clip-Machern inszenierte, das heißt also künstliche visuelle Ebene ersetzt. Beim Spielfilm fungiert in der Regel die Musik als Ergänzung zur Bild- und Handlungsfolge, beim Videoclip dagegen eröffnet die optische Ebene zur Musik den Einstieg in illustrative, situative, narrative und/oder assoziative Bilderwelten.

### 3. Filmmusik- und Videoclip-Analyse

Zwangsläufig stellt sich jetzt die Frage, ob die hier skizzierten Hinweise und Überlegungen zur Konvergenz der Sinne und zur audiovisuellen Musikwahrnehmung Konsequenzen für das Verständnis und die Analyse von Filmmusik und Videoclip haben. Betrachtet man Musikanalyse als ein Hilfsmittel, bei dem man – so Herbert Bruhn in *Musikwissenschaft. Ein Grundkurs* (1998: 493) – ein »Musikstück in seine erklingenden Bestandteile zerlegt und sie dann nach verschiedenen Kriterien beurteilt und zueinander in Beziehung setzt«, dann ergeben sich daraus eine Fülle von Fragestellungen. Sie führen von dem musikalischen Objekt und seiner handwerklichen Machart über Produktionsbedingungen und Aspekte des gesellschaftlichen Verwendungszusammenhangs bis hin zur Einbeziehung des Hörenden in den Analyseprozess. Denn schließlich wendet sich Musik letztlich immer an ihn. Seine Bewertung entscheidet über Erfolg oder Nichterfolg der musikalischen Produkte. Genau in diesem Sinn hat die bisherige Geschichte der Populärmusikforschung mit ihren intensiven Diskussionen auf dem Weg zu sinnvollen Analysemethoden zumindest eines deutlich werden lassen: Jede ernsthafte wissenschaftliche Auseinandersetzung mit Musik ist in einem multidimensionalen Feld von Beziehungen zu verorten, das möglichst umfassend alle Variablen von der Musikproduktion bis zu ihrer Rezeption erfassen sollte (Rösing 2002).

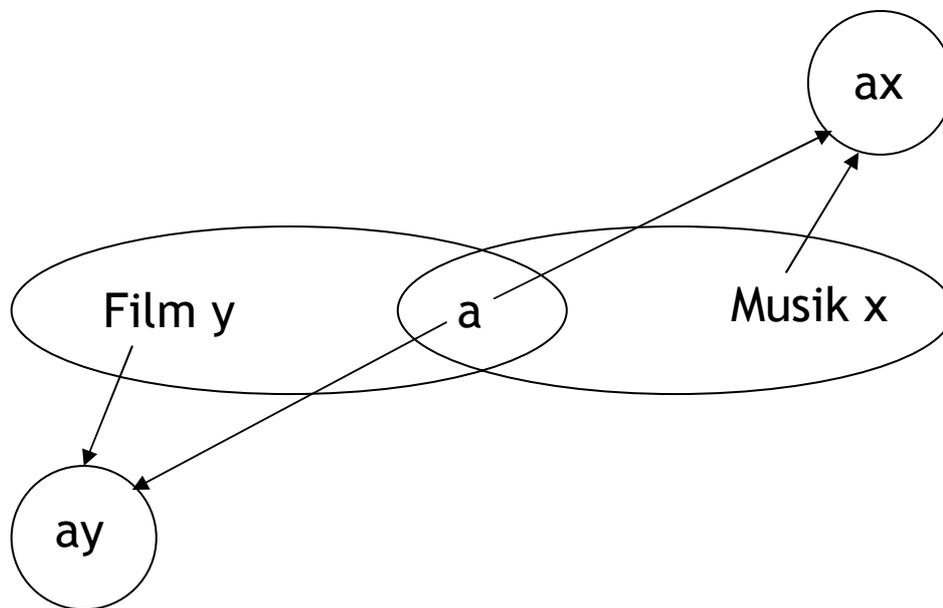
Doch das ist leichter gesagt als getan. In den Arbeiten über Filmmusik z.B. wird besondere Aufmerksamkeit vor allem der Frage nach den Funktionen bzw. Wirkungsweisen von Musik im Film entgegengebracht. Eine Vielzahl von diesbezüglichen Kategorisierungsversuchen mag als Ansatzpunkt von Analyse durchaus hilfreich sein – vorausgesetzt, man hat sich erst einmal für eine der Funktionssystematiken entschieden: etwa die jüngst von Claudia Bullerjahn (2001) vorgelegte mit der Unterscheidung in Metafunktionen, Funktionen im engeren Sinn und filmmusikalische Techniken. Eine nach dieser Vorgabe von André Matthias (2001) durchgeführte Detailanalyse der Musik zum Film *Braveheart* kommt allerdings zu dem Fazit, dass gemäß den Bullerjahnschen Kategorisierungsvorschlägen durchaus gleichzeitig »alle Bereiche von den dramaturgischen und epischen über die strukturellen bis hin zu den persuasiven Funktionen [Funktionen im engeren Sinn] abgedeckt werden«. Selbst die verschiedenen angewendeten filmmusikalischen Techniken – deskriptive Technik, Mood-Technik, Leitmotiv-Technik (es fehlt lediglich die von Bullerjahn zusätzlich genannte Baukasten-Technik) – erfüllen hier letztlich »alle Funktionen mit allen Mitteln« (Matthias 2001: 150).

Auch neueste Detailanalysen von Videoclips weisen in diese Richtung. Die von Winfried Pape und Kai Thomsen (1997) im Überblick zusammengestellten Kategorisierungsangebote – etwa nach den Kriterien: reine Performance, »Konzept«-Performance, narratives, situatives, illustratives Konzept sowie verschiedene filmische Mittel und Schnitttechniken – führten Sonja Henscher (2001) bei der vergleichenden Analyse der Darstellung afro-amerikanischer Rapperinnen im Videoclip zu dem wenig überraschenden Ergebnis, dass die gängigen Einteilungen in Clip-Typen und die hinlänglich bekannten stereotypen Aussagen zur Präsentation »schwarzer« Frauen im Videoclip auf Wahrnehmungskonstruktionen beruhen, die der Vielfalt der tatsächlich nachweisbaren Präsentationsformen nicht so ohne weiteres gerecht werden. Derartige Analyseergebnisse verweisen darauf, dass (1) die audiovisuelle Objektebene eine sehr große Informationsdichte hat. Diese zwingt im Rezeptions- und Interpretationsprozess zur Auswahl, und zwar gerade dann, wenn die »optische und akustische Aufmerksamkeit gleichermaßen erregt« aktiviert wird, wie das Klaus Neumann-Braun und Axel Schmidt in ihrem Buch *Viva MTV! Popmusik im Fernsehen* (1999: 10) unter Absage an ein angeblich optisches Zeitalter formulieren. Und sie verweisen (2) auf die informationsschaffende Kraft des menschlichen Gehirns, das auf der Basis intermodaler Wahrnehmung Beziehungen zwischen der Text-, Bild- und Musikebene herstellt.

Wie das funktionieren kann, hat Nicolas Cook in seinem Buch *Analyzing Musical Multimedia* (1998) mit dem »Congruence-associationist-model« zu beschreiben versucht. Demnach überschneiden sich die Bedeutungsebenen von bewegten Bildern und Musik im Film zwar zum Teil, führen aber gemäß der Konvergenz der Sinne und der beziehenden Wahrnehmung auch zur Konstruktion von je nach der Rezeptionssituation immer wieder neuen Bedeutungszusammenhängen. Je nach Art und Schwerpunktsetzung der Aufmerksamkeitszuwendung gegenüber dem audiovisuellen Gesamtgeschehen ändert sich die Wahrnehmung. So kann z.B. (vgl. Graphik 1) durch die Musik eine Aufmerksamkeitsfokussierung erfolgen, die dazu führt, dass nur ein bestimmter Ausschnitt (a) aus dem gesamten visuellen Informationsangebot (y) auch tatsächlich und bewusst rezipiert wird. Die Deutung des visuellen Wahrnehmungsausschnitts erfolgt hier in Abhängigkeit von der gleichzeitig erklingenden Musik (x). Mit anderen Worten: Die Musik beeinflusst die Wahrnehmung des visuellen Geschehens durch assoziative Verknüpfung (ax). Doch auch der umgekehrte Fall ist denkbar. In Bezug auf das bewusst rezipierte visuelle Geschehen (y) wird nur ein Teil (a) von der Musik und ihren Strukturen (x) wahrgenommen und in Bezug zum Bild interpretiert. Dann beeinflusst das filmische Geschehen die Bedeutung der Musik (ay).

Graphik 1: Wahrnehmungsmodell nach Nicolas Cook (1998).

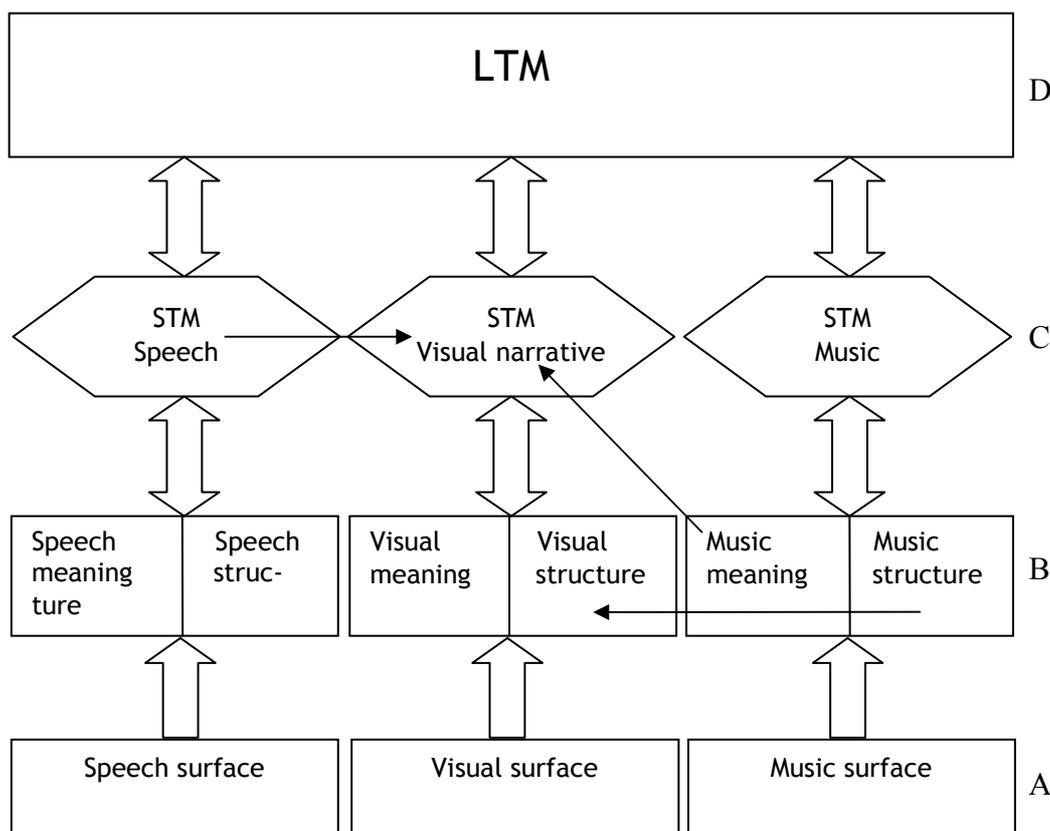
Je nach Aufmerksamkeitszuwendung beeinflusst Musik die Filmwahrnehmung oder aber das filmische Geschehen die Musikwahrnehmung.



Natürlich ist die rezeptive Wirklichkeit weit komplizierter als es das Modell beschreibt, das auf die vielen möglichen Zwischenformen und im Zeitverlauf sich ständig verändernden Aufmerksamkeitsfokussierungen nicht weiter eingeht. Dazu hat Annabel J. Cohen in dem von Patrik N. Juslin und John A. Sloboda herausgegebenen Sammelband über *Music and Emotion* (2001) ein differenzierteres Wahrnehmungsmodell beschrieben (vgl. Graphik 2). Es berücksichtigt die drei Informationskanäle Sprache, Bild, Musik mit ihren objektiv gegebenen Reizstrukturen (Ebene A) und zeigt deren mentale Verarbeitung auf den Ebenen B bis D an. Struktur und Bedeutung der sprachlichen, visuellen und musikalischen Information (Ebene B) werden im Kurzzeitgedächtnis (STM – short term memory) zu Verarbeitungseinheiten zusammengefasst. Diese sind – über den Vorgang der Aufmerksamkeitslenkung und der assoziativen Zuordnung – durch die im Langzeitgedächtnis (LTM – long term memory) abgespeicherten visuellen und auditiven Erfahrungsinventare (Ebene D) ebenso mitgeprägt wie durch die tatsächliche Reizstruktur der sprachlichen, optischen und musikalischen Information. Auf den Ebenen C und B kommt es also, gemäß den zuvor dargelegten neurophysiologischen Gegebenheiten, zu einem Austausch von bottom-up- und top-down-Prozessen, allerdings in unterschiedlicher Gewichtung (siehe Pfeilstärke). Die Auswirkung einer möglichen primär musikbezogenen Aufmerksamkeitslenkung für die Wahrnehmung von Sprache und Bild in Analogie zum

ax-Wahrnehmungsmodus im Cookschen Modell ist durch die schwarzen Pfeile angedeutet.

Graphik 2: Multimediales Wahrnehmungsmodell nach Annabel J. Cohen (2001).



Aus den Ergebnissen der bisher vorliegenden Detailanalysen von Filmmusik und Videoclip und den zwei hier vorgestellten rezeptionspsychologischen Wahrnehmungsmodellen ergeben sich Konsequenzen für zukünftige Film-musik- und Videoclipuntersuchungen, die hier abschließend angedeutet seien. Die Struktur multimedialer Produkte ist durch eine Informationsdichte gekennzeichnet, die beim Rezipieren zu einer wie auch immer gearteten Auswahl durch Aufmerksamkeitsfokussierung zwingt. Theodor W. Adorno und Hanns Eisler (1969: 117) waren darum der Ansicht, »daß jeder Mensch beim Hören von Musik [zum Film] andere [...] und nur für sich selbst bedeutsame Assoziationen besitzt«. Dem kann, bezogen auf die Objektseite, allein mit dem Bemühen um eine minutiöse, detailgenaue Deskription der strukturellen Bild-, Sprach- und Musikebene begegnet werden. Alle darauf gründenden Auswertungen und Interpretationen dürfen jedoch nicht verabsolutiert werden. Hier kann es sich immer nur um Interpretationsangebote handeln, die, wie Heinrich Bessler für die Musikrezeption schon 1959 bün-

dig belegt hat, durchaus auch zeitgeschichtlich bedingt sein mögen. Verschiedene neuere Arbeiten zur aktiven Konstruktionsleistung beim Musikhören (Rösing 2000) und zur musikalischen Bedeutungsambivalenz im Sinn einer Projektionsfläche für inhaltliche Zuschreibungen (Rösing 2002) verweisen darauf ebenso wie z.B. der Artikel »Musikverstehen – ein Mißverständnis?« von Klaus-Ernst Behne. Der Verstehensbegriff sei, so sein Resümee, in Bezug auf das nichtdiskursive Medium Musik »in höchstem Maße fragwürdig« (Behne 1994: 181).

Die interpretatorische Unschärferelation nimmt bei artifiziellen Produkten zwangsweise um ein Vielfaches zu, wenn Musik nicht alleine, sondern zusammen mit Sprache und bewegten Bildern erklingt. Mit der daraus resultierenden Informationszunahme bzw. größeren Informationsdichte pro Zeiteinheit, mit dem Zusammenspiel einer Fülle von symbolischen Codes und assoziativen Semantemen werden die Interpretationsfreiräume größer und nicht, wie so oft behauptet wird, kleiner – selbst noch dort, wo es primär um Bedeutungsergänzung und weniger um kommentierende, kontrastierende, Emotionen auslotende usw. Informationen geht. Jede Art der Bedeutungsergänzung im multimedialen Zusammenspiel bietet Stoff für neue Formen der beziehenden Wahrnehmung und damit auch der Ausdeutung. Je mehr Dimensionen ein Wahrnehmungsangebot enthält, umso vielfältiger sind die Assoziations- und Verknüpfungsmöglichkeiten unter Berücksichtigung der jeweils gegebenen intra- wie intersubjektiven Erfahrungsinventare. Oder, mehr metaphorisch gesagt, umso reichhaltiger wird die Bilderwelt der Klänge und die Klangwelt der Sinne. Ich bin darum der Ansicht, dass Filmmusik- und Videoclip-Analysen um so mehr an Erkenntnisgewinn versprechen, je genauer (1) die Analysemodelle und Beschreibungskriterien in fachübergreifender Zusammenarbeit von Vertretern und Vertreterinnen der entsprechenden Disziplinen ausgearbeitet werden, je genauer aber auch (2) zwischen der Beschreibung der strukturellen Objektebene als Reizauslöser für Wahrnehmungsprozesse und der mentalen Verarbeitung als aktiver Konstruktions- bzw. Interpretationsleistung unterschieden wird, sowie (3) das jeweilige Erkenntnisinteresse im Sinn einer forschungsleitenden Fragestellung formuliert wird. Das Erkenntnisinteresse ist dann eben nicht auf das Ganze, sondern nur auf einen Bedeutungsausschnitt, eine von vielen möglichen Perspektiven der Betrachtung gerichtet. Der Vorteil besteht darin, dass eine begrenzte Fragestellung durch solide Deskription sowie empirische Überprüfung verifiziert bzw. falsifiziert werden kann. Das Ergebnis ist ein Interpretationsangebot, das in Bezug auf die jeweilige Fragestellung in einer bestimmten soziokulturellen Situation und bei Personen mit einer jeweils spezifischen Sozialisation seine Bedeutung und damit Gültigkeit hat. Darü-

ber hinausgehende Wahrheiten sind im Forschungsbereich der Kulturwissenschaften nicht zu haben.

## Literatur

- Adorno, Theodor W. / Eisler, Hanns (1969). *Komposition für den Film*. München: Rogner & Bernhard.
- Adorno, Theodor W. (1976). *Anweisungen zum Hören neuer Musik* (= Gesammelte Schriften 15). Frankfurt/M.: Suhrkamp, S. 188-248.
- Altenmüller, Eckart (2000). »Neue Querverbindungen durch musikalisches Lernen – oder warum es im Gehirn kein festes Musikzentrum gibt.« In: *Querverbindungen*. Hg. v. Georg Mantel. Mainz: Schott, S. 90-97.
- Arnheim, Rudolf (1974). *Film als Kunst*. München: Hanser (Originalausgabe 1932).
- Behne, Klaus-Ernst (Hg.) (1987). *Film – Musik – Video oder die Konkurrenz von Auge und Ohr*. Regensburg: Bosse.
- Behne, Klaus-Ernst (1994). »Musikverstehen – ein Mißverständnis?« In: *Gehört – Gedacht – Gesehen. Zehn Aufsätze zum visuellen, kreativen und theoretischen Umgang mit Musik*. Hg. v. Klaus-Ernst Behne. Regensburg: Con Brio, S. 167-183.
- Bessler, Heinrich (1959). *Das musikalische Hören der Neuzeit* (= Bericht über die Verhandlungen der Sächs. Akademie der Wissenschaften Leipzig, Phil.-hist. Klasse Bd. 104, Heft 6). Berlin.
- Blaukopf, Kurt (1989). *Beethovens Erben in der Mediamorphose. Kultur- und Medienpolitik für die elektronische Ära*. Heiden (CH): Niggli.
- Bruhn, Herbert (1998). »Analyse.« In: *Musikwissenschaft. Ein Grundkurs*. Hg. v. Herbert Bruhn und Helmut Rösing. Reinbek: Rowohlt, S. 493-507.
- Bullerjahn, Claudia (2001). *Grundlagen der Wirkung von Filmmusik*. Augsburg. Wißner.
- Cohen, Annabel J. (2001). »Music as a Source of Emotion in Film.« In: *Music and Emotion. Theory and Research*. Hg. v. Patrik N. Juslin und John A. Sloboda. Oxford: Oxford University Press, S. 249-272.
- Cook, Nicolas (1998). *Analysing Musical Multimedia*. Oxford: Clarendon Press.
- EGGEBRECHT, Hans Heinrich (1994). *Zur Geschichte der Beethoven-Rezeption*. Laaber: Laaber.
- Hausheer, Cecilia / Schönholzer, Annette (Hg.) (1994). *Visueller Sound. Musikvideos zwischen Avantgarde und Populärkultur*. Luzern: Zyklus.
- Henscher, Sonja (2001). *Afro-amerikanische Rapperinnen im Videoclip. Eine vergleichende Analyse von »Doo Wop (That Thing)« (Lauryn Hill), »No matter what they say« (Lil' Kim) und »She's a Bitch« (Missy Elliott)*. Hamburg (unveröff. Mag.-Arb., Musikwissenschaftliches Institut der Universität).
- Hornbostel, Erich Moritz v. (1924). »Die Einheit der Sinne.« In: *Melos* 4, S. 290-297.
- Huber, Kurt (1923). *Der Ausdruck musikalischer Elementarmotive. Eine experimentallypsychologische Untersuchung*. Leipzig.
- Hurte, Michael (1982). *Musik, Bild, Bewegung. Theorie und Praxis auditiv-visueller Konvergenzen*. Bonn: Verlag für Systematische Musikwissenschaft.
- Imberty, Michel (1971). »Perception et significations de œuvre musicale.« In: *Psychologie Française* 16, S. 137-155 (Übersetzung in: Helmut Rösing [Hg.] [1983]. *Rezeptionsforschung in der Musikwissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft, S. 295-321).

- Jewanski, Jörg (1999). *Ist C = Rot? Eine Kultur- und Wissenschaftsgeschichte zum Problem der wechselseitigen Beziehung zwischen Ton und Farbe. Von Aristoteles bis Goethe* (= Berliner Musikstudien 17). Sinzig: Studio.
- Kant, Immanuel (1799). *Kritik der Urteilskraft*. Berlin.
- Kircher, Athanasius (1650). *Musurgia universalis*. Rom (Nachdr. Hildesheim: Olms 1970).
- Klauwell, Otto (1968). *Geschichte der Programmmusik von ihren Anfängen bis zur Gegenwart*. Wiesbaden: Sändig (Erstveröffentl. 1910).
- Kleinen, Günter (1994). *Die psychologische Wirklichkeit der Musik. Wahrnehmung und Deutung im Alltag*. Kassel: Bärenreiter.
- Kneif, Tibor (1970). »Ideen zu einer dualistischen Musikästhetik.« In: *International Review of Music, Aesthetics and Sociology* 1, S. 15-34.
- Kracauer, Siegfried (1964). *Theorie des Films. Die Errettung der äußeren Wirklichkeit* (= Schriften 3, hg. v. Karsten Witte). Frankfurt/M.: Suhrkamp.
- Marks, Lawrence E. (1978). *The Unity of Senses: Interrelations among the Modalities*. New York: Academic Press.
- Matthias, André (2001). *Die Musik zum Film ›Braveheart‹. Bestandsaufnahme und Analyse*. Hamburg (unveröff. Mag.-Arb., Musikwissenschaftliches Institut der Universität).
- Neumann-Braun, Klaus / Schmidt, Axel (Hg.) (1999). *Viva MTV! Popmusik im Fernsehen*. Frankfurt/M.: Suhrkamp.
- Pape, Winfried / Thomsen, Kai (1997). »Zur Problematik der Analyse von Videoclips.« In: *Step across the Border: Neue musikalische Trends – neue massenmediale Kontexte*. Hg. v. Helmut Rösing (= Beiträge zur Populärmusikforschung 19/20). Karben: CODA, S. 200-226.
- Pech, Karel (1969). *Hören im optischen Zeitalter*. Karlsruhe: Braun.
- Rösing, Helmut (Hg.) (1993). *Spektakel/Happening/Performance. Rockmusik als ›Gesamtkunstwerk‹*. Mainz: Villa Musica.
- Rösing, Helmut (1997). »Musikalische Sozialisation und Musikpädagogik.« In: *Persönlichkeitsentfaltung durch Musikerziehung*. Hg. v. Josef Scheidegger und Hubert Eiholzer. Luzern: Nepomuk, S. 164-183.
- Rösing, Helmut (1998a). »Synästhesie.« In: *Die Musik in Geschichte und Gegenwart*. Sachteil Bd. 9. Hg. v. Ludwig Finscher. Kassel-Stuttgart: Bärenreiter-Metzler (2. Aufl.), Sp. 168-185.
- Rösing, Helmut (1998b). »Musik – ein audiovisuelles Medium. Über die optische Komponente der Musikwahrnehmung.« In: *Musikwissenschaft zwischen Kunst, Ästhetik und Experiment*. Festschrift Helga de la Motte-Haber zum 60. Geburtstag. Hg. von Reinhard Kopiez u.a. Würzburg: Königshausen und Neumann, S. 451-463.
- Rösing, Helmut (2000). »Zur medialen Konstruktion musikalischer Lebenswelten. Eine kritische Bestandsaufnahme.« In: *Populäre Musik im kulturwissenschaftlichen Diskurs*. Hg. v. Helmut Rösing und Thomas Phleps (= Beiträge zur Populärmusikforschung 25/26). Karben: CODA, S. 11-23.
- Rösing, Helmut (2002). »Männlichkeitssymbole in der Musik – eine Spurenlese.« In: *Festschrift Ekkehard Jost zum 65. Geburtstag*. Hg. v. Bernd Hoffmann, Franz Kerschbaumer, Franz Krieger und Thomas Phleps (= Jazzforschung / Jazz Research 34). Graz: Akademische Druck- und Verlagsanstalt, S. 243-256.
- Roth, Gerhard (1997). *Das Gehirn und seine Wirklichkeit. Kognitive Neurobiologie und ihre philosophischen Konsequenzen*. Frankfurt/M.: Suhrkamp.
- Rouget, Gilbert (1985). *Music and Trance: A Theory of Relations between Music and Possession*. Chicago/London.

- Salmen, Walter (1988). *Das Konzert. Eine Kulturgeschichte*. München.
- Schönhammer, Rainer (1988). *Der Walkman. Eine phänomenologische Untersuchung*. München (Phil. Diss.).
- Schneider, Norbert Jürgen (1997). *Komponieren für Film und Fernsehen. Ein Handbuch*. Mainz: Schott.
- Wellek, Albert (1928). *Doppelempfinden und Programmmusik. Beiträge zur Psychologie, Kritik und Geschichte der Sinnenentsprechung und Sinnensymbolik*. Wien (Phil. Diss., Musikwissenschaftliches Institut der Universität).
- Werner, Heinz (1966). »Intermodale Qualitäten (Synästhesien).« In: *Handbuch der Psychologie*. Bd. I, 1: Wahrnehmung und Bewußtsein. Hg. v. Wolfgang Metzger und Heiner Erke. Göttingen: Hogrefe, S. 278-303.

## Abstract

This paper develops a theory of audiovisual music perception on the basis of neurophysiological and psychological research on the intermodality of perception and convergence of the senses. Six modes of perception are described (from music perception to visual perception), which are then discussed regarding the analysis of music in film and video. It is shown that each different way of adding meaning to the multimedial interplay of music, picture and text produces new material for new forms of intermodal perception. The more dimensions an object of perception contains the greater is the variety of possible interpretations, if all intra- as well as intersubjective inventories of experience are taken into consideration. An analysis that remains immanent to the structure of film music or music video can therefore result only in suggestions for interpretation, which require empirical tests regarding specific queries and specific groups of audiences.