

Forschungsberichte

Jörg Langner & Reinhard Kopiez

Entwurf einer neuen Methode der Performanceanalyse auf Grundlage einer Theorie oszillierender Systeme (TOS)

Einleitung

In den letzten 15 Jahren ist die Performanceforschung zu einem zentralen Gegenstand der Musikpsychologie geworden. Im Mittelpunkt ihres Interesses stehen der Spieler und seine Tätigkeit im allgemeinen und der durch den Spieler gestaltete ausdrucksvolle Vortrag eines Musikstückes im besonderen. Der Begriff der Performance umfaßt dabei sowohl allgemeine kognitionspsychologische Aspekte, wie z.B. Lernen, Gedächtnis oder Wahrnehmung, als auch den speziellen Aspekt des musikalischen Ausdrucks. Die Bedingungsvariablen des musikalischen Ausdrucks beim Spieler zu bestimmen, wurde zu einem Hauptanliegen der Performanceforschung. Prinzipiell bieten sich hierzu zwei Verfahren an: erstens die Methode Analyse-durch-Messung (Analysis-by-Measurement) und zweitens die Methode der Analyse-durch-Synthese (Analysis-by-Synthesis). Das erste Verfahren der Analyse-durch-Messung hat die längere Forschungstradition und kann zumindest bis auf die Studien von Seashore (1936, 1938) zurückgeführt werden. Eine Reaktivierung dieser Methode fand durch die technologisch innovativen Arbeiten von Shaffer (1980, 1981), Gabrielsson (1987) und in jüngster Zeit Repp (1990, 1992) statt. Allen Autoren gemeinsam ist, daß sie die vom Spieler erzeugten Performancedaten als Ausgangspunkt ihrer Analysen verwenden. Hierbei werden sowohl MIDI- als auch akustische Daten (von Schallplatten) verwendet. Das Forschungsparadigma dieses Ansatzes ist die Erstellung einer sogenannten Timingkurve, die agogische Schwankungen, das sogenannte "expressive timing", wiedergibt: Wie z.B. aus der Analyse des Themenverlaufes von Mozarts A-Dur-Sonate KV 331 in Ga-

brielsson (1987) ersichtlich, wird der musikalische Ausdruck des Spielers – unter fast ausschließlicher Berücksichtigung des Timings – als Abweichung der realen Tondauern von einem gedachten nominellen, metronomisch exakten Wert dargestellt. Die Ergebnisgrafik zeigt dann einen mehr oder weniger deutlich ausgeprägten zickzackförmigen Verlauf.

Das zweite Verfahren, die Tätigkeit des Spielers zu analysieren, besteht in der Simulation einer menschlichen Performance: Aus dem Spiel hervorragender Musiker werden intuitiv Regeln über den Zusammenhang zwischen dem Notentext und der erklingenden Realisation abgeleitet, und in einem zweiten Schritt erfolgt durch Implementierung dieser Regeln in einen Algorithmus mit anschließender Applikation auf gespeicherte Melodien eine Ausdruckssynthese. Diese Methode wurde maßgeblich durch Johan Sundberg von der Musikakademie Stockholm entwickelt (vgl. den zusammenfassenden Überblick zu dieser Methode in Sundberg, 1988, 1991). Die Methode besteht darin, aus einer überzeugenden Performancesimulation und den dabei beobachteten Zusammenhängen auf Prozesse zu schließen, die im Denken und Spielen eines Musikers ablaufen. Der Analyse-durch-Synthese-Ansatz hat in den letzten Jahren durch die Verwendung leistungsfähiger Computer und neuer Softwarekonzeptionen große Fortschritte gemacht. Hervorzuheben ist hier vor allem die Steuerung der Regelapplikation durch eine Programmarchitektur, die auf einem Netzwerkmodell beruht (vgl. Battel, Bresin, De Poli & Vidolin, 1993). Zusätzlich gibt es einen Neuansatz in der Syntheseforschung durch Mazzola (1993), der die »Logik der musikalischen Performance« untersucht, ohne auf die Trial-and-Error-Methode des Analyse-durch-Synthese-Ansatzes zurückzugreifen (eine ausführliche Darstellung dieser Methoden und ihrer Unterschiede findet sich in Kopiez, 1995).

Zwar ist momentan noch keine Prognose darüber möglich, welche der beiden Methoden sich als die erfolgreichere erweisen wird, doch liegt unserer Arbeit die Überzeugung zugrunde, daß die Interpretationen hervorragender Musiker durch die bestehenden Verfahren des Analyse-durch-Messung-Ansatzes nur oberflächlich ausgewertet werden. Analysen auf Basis der Timing-Kurve schöpfen die in den Performancedaten enthaltenen Informationen nicht aus. Dieses Manko hat seine Ursache in den bisher verwendeten Verfahren, nicht jedoch in der Methode der Datenanalyse selbst.

Am deutlichsten läßt sich der Ansatzpunkt unserer Vorgehensweise durch eine Kritik bestehender Methoden und einem daraus resultierenden Forderungskatalog formulieren:

1. Wir postulieren, daß die bisher verwendeten Timingkurven ein unzureichendes Bild einer Performance wiedergeben. Diese Kurven gehen von einer lediglich eindimensionalen Zeitstruktur aus. Wie neuere theoretische Arbeiten dagegen aufzeigen konnten, ist die musikalische Zeitgestaltung jedoch stets mehrdimensional: Zeitgestaltung findet auf unterschiedlichen Dauernebenen statt, wobei das Tempo auf den Ebenen bestimmter Notenwerte, ganzer Takte oder Taktgruppen bis hin zu den übergeordneten Schichten ganzer Formteile variabel sein kann. Deshalb ist es unumgänglich, die monolineare Vorstellung von Tempo, auf der die Arbeit mit einer einzigen Tempokurve pro Stück beruht, aufzugeben und stattdessen von Tempohierarchien oder der Zeitarchitektur einer Performance auszugehen (vgl. zu diesen Begriffen Mazzola, 1993; Mazzola & Zahorka, 1994; Gottschewski, 1993).

Forderung: Eine Analysemethode muß die volle Komplexität der Zeitgestaltung einer Performance wiedergeben können.

2. In der forschungsgeschichtlichen Entwicklung der Analyse-durch-Messung-Verfahren ist ein zunehmender Reduktionismus zu beobachten: Untersuchten Hartmann (1932) oder Seashore (1938) noch die Zeitverläufe ganzer Sonatensätze oder zumindest mehrminütiger Abschnitte daraus, so werden bei Gabrielsson (1987) nur noch die ersten acht Takte des Themas aus Mozarts A-Dur-Sonate KV 331 analysiert. Repp (1992) mißt gar nur noch die Dauern der ersten sechs Noten der melodischen Anfangsfigur aus Schumanns Träumerei. Wir knüpfen mit unserer Methode bewußt an die ältere Forschungstradition an und wenden uns gegen reduktionistische Ansätze.

Forderung: Eine adäquate Analysemethode muß dazu in der Lage sein, sowohl die Mikro- als auch die Makrozeitverläufe eines Stückes darzustellen. Da sich die volle Zeitkomplexität einer Performance nur bei Betrachtung des gesamten Stückes zeigen kann, muß die Analyseperspektive auch die Dauer eines ganzen Stückes bzw. Satzes umfassen.

Methoden

Eine Antwort auf die gestellten Forderungen bietet die Theorie oszillierender Systeme (TOS), wie sie gegenwärtig von Langner (i.V.) entwickelt wird, vorläufig noch auf den Bereich des Rhythmus beschränkt. Diese Theorie beruht auf der modellhaften Annahme, daß jeder musikalische Verlauf in der menschlichen Wahrnehmung eine für diesen Verlauf charakteristische

Konstellation von Oszillationen auslöst (dies sind Schwingungen oder schwingungsähnliche periodische Vorgänge) und auf dieser Basis zahlreiche Phänomene des musikalischen, insbesondere des rhythmischen Erlebens erklärt werden können.

Der Grundgedanke, die Wahrnehmung von Rhythmus im Zusammenhang mit periodischen Prozessen zu sehen, ist in der musikpsychologischen Literatur vielfach anzutreffen. Der Begriff der Oszillation selbst findet sich etwa bei Gjerdingen (1993) oder Large & Kolen (1994), aber auch zahlreiche weitere theoretische Ansätze enthalten die Annahme schwingungsähnlicher Vorgänge explizit oder implizit. Erwähnt seien hier die »rhythmograms« in den Publikationen von Todd (1994a, b), die »pulse sensations« bei Parncutt (1994), die »rhythmic levels« in einer Arbeit von Rosenthal (1992) sowie das »hierarchical clock model« von Povel & Essens (1985). Auch in anderen Zusammenhängen, zum Beispiel verbunden mit dem Phänomen der Aufmerksamkeit, finden wir die Annahme periodischer Vorgänge, so etwa bei Jones & Yee (1993).

Die meisten der mit schwingungsähnlichen Prozessen arbeitenden theoretischen Ansätze beschränken sich auf einige wenige Oszillatoren, wie sie etwa den verschiedenen metrischen Ebenen in einem 4/4-Takt zugeordnet werden können (und postulieren dann periodische Vorgänge z.B. auf Viertel-, Halb- und Ganznotenebene). Eine solche Beschränkung führt sogleich zu Schwierigkeiten, wendet man sich von der Analyse eines Notentextes ab und legt Daten zugrunde, die bei einer realen Performance entstehen: Die »expressive deviations« des Musikers sind mit der Existenz von einigen wenigen, mit konstanter Periode oszillierenden Einheiten nicht verträglich und verursachen etwa einer Computer-Analyse von Performance-Daten erhebliche Probleme (siehe hierzu Desain & Honing, 1994).

Grundsätzlich gibt es zwei Möglichkeiten, auf diese Schwierigkeit zu reagieren:

1. Die Oszillatoren selbst werden mit der Fähigkeit einer flexiblen Anpassung an das Tempo ausgestattet. Large (1994) sowie Large & Kohlen (1994) realisieren diese Option mit einem Computermodell.
2. Die Anzahl der mit fester Frequenz schwingenden Einheiten wird erhöht, dies ermöglicht ein dichteres Frequenzspektrum und somit ein flexibleres System. Diesen Ansatz finden wir bei Todd (1994a, b), allerdings bislang nicht konsequent durchgeführt.

Mit dem der vorliegenden Studie zugrunde liegenden Modell wird der zweite Weg beschritten: Es basiert auf einem Set von 145 Oszillatoren mit

jeweils fester Frequenz, verteilt auf einen Bereich zwischen 0,002 und 24 Hz. Er umfaßt somit insbesondere alle musikalischen Tempi, die Frequenz von 2 Hz zum Beispiel entspricht einem Tempo von M.M. = 120 (die Metronomzahlen ergeben sich aus den Hz-Werten per Multiplikation mit 60). Die Grenzen dieses Spektrums liegen allerdings um einiges weiter, als sie in der Literatur für »rhythmische Frequenzen« angegeben werden (0,3 bis 20 Hz zum Beispiel bei Gjerdingen, 1993, S. 503; 0,02 bis 20 Hz bei Todd, 1994a, S. 395).

Erklärungsbedarf besteht für die zunächst außerordentlich niedrig erscheinende Untergrenze (der Frequenz von 0,002 Hz entspricht eine Schwingungsdauer von über acht Minuten!): Eine zentrale Hypothese bei Langner (i.V.) besagt, daß die Aktivierung von sehr niederfrequenten Oszillatoren a) mit der Wahrnehmung von großräumigen musikalischen Formzusammenhängen korrespondiert und b) mit positiven ästhetischen Bewertungen von Komposition und Performance korreliert. Folglich muß das Modell Oszillatoren enthalten, deren Schwingungsdauer großräumigen Zusammenhängen entspricht. Daß es Menschen gibt, die eine musikalisch gefüllte Zeitstrecke von acht Minuten als zusammenhängend erleben können, erscheint unstreitig, die Untergrenze bei 0,002 Hz von daher sicher nicht zu niedrig angesetzt. Für eine präzise Bestimmung wären spezielle Studien erforderlich. Erwähnt sei, daß die Vermutung eines Zusammenhanges solch niederfrequenter Oszillationen mit der Formwahrnehmung vereinzelt in der Literatur begegnet (z.B. bei Parncutt, 1994, S. 437).

Die Aufteilung der 145 Oszillatoren auf den Frequenzbereich erfolgt nicht linear, sondern näherungsweise logarithmisch (d.h. sie liegen im unteren Bereich wesentlich dichter als im oberen); dies entspricht sowohl der musikalischen Erfahrung mit Tempo (auch das Metronom ist ungefähr logarithmisch gestuft), als auch den Resultaten empirischer Studien (Epstein, 1985). Die Anzahl ist mit 145 so gewählt, daß sich in weiten Bereichen des Spektrums zwei benachbarte Einheiten in ihrer Frequenz um ca. 4% unterscheiden; dies korrespondiert abermals mit den Schrittweiten auf dem Metronom und in der Größenordnung mit experimentellen Ergebnissen zur Genauigkeit von Tempowahrnehmung (Drake, Botte & Baruch, 1992).

Als Input für die Berechnungen dient der Lautstärkeverlauf des zu untersuchenden Musikstückes. Dieser kann entweder aus den Noten abgeleitet oder einer realen Performance entnommen werden (siehe Abbildung 1). In diesem Fall wird der Schallpegelverlauf einer Aufnahme in Schritten von 5 oder 10 ms gesampelt und in ein Daten-File umgewandelt. (Unsere Beispie-

le wurden mit dem Programm Energy der Audioanalysesoftware CARL [California Acoustic Research Laboratories] gesampelt). Erste Berechnungen mit solchen Daten zeigten, daß eine Umwandlung der dB-Werte (Schallpegel) auf die Sone-Skala (subjektive Lautstärke, Lautheit) zu musikalisch noch plausibleren Ergebnissen führt. Weitere Feinheiten der Lautstärkewahrnehmung einzubeziehen (Frequenzabhängigkeit, Maskierung), erscheint derzeit nicht erforderlich, ist jedoch grundsätzlich möglich.

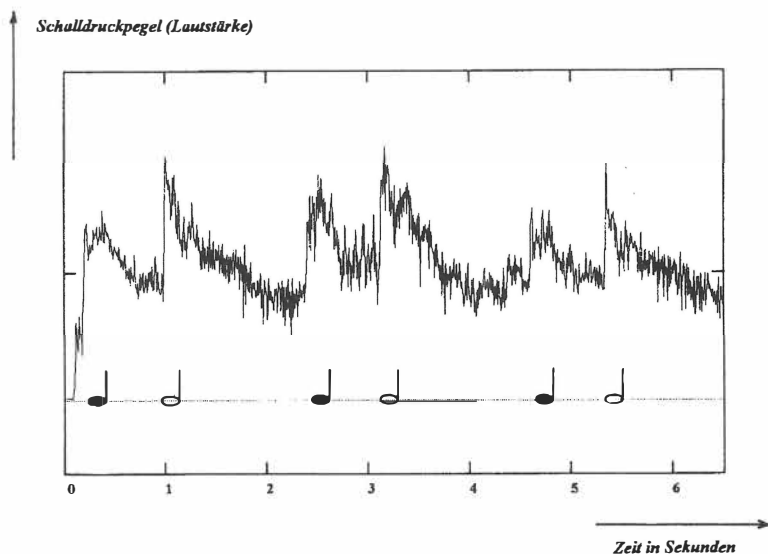


Abb. 1 Mit einer Rate von 5 ms gesampelter Schalldruckpegelverlauf, wie er sich durch das Legatospiele eines Pianisten ergibt. Die Notenwerte der gespielten Tonfolge sind an entsprechender Stelle auf der Zeitachse eingetragen. Ein solcher Pegelverlauf dient als Input für die hier verwendeten Algorithmen.

Von zentraler Bedeutung ist die Frage, nach welchen Gesetzmäßigkeiten die Anregung der Oszillationen durch den musikalischen Input erfolgen soll. Den Ausschlag hierbei geben letztlich die Anforderungen, welche an das Modell zu stellen sind, sowie der angestrebte Grad an neurophysiologischer Plausibilität (ein Oszillator kann als Population periodisch feuerner Nervenzellen gedacht werden). Zu fordern ist in jedem Fall eine getreue Abbildung des musikalischen Tempos; beispielsweise erwarten wir, daß eine Folge von gleichlangen Noten im Tempo M.M. = 60 ein Aktivierungsmaximum bei 1 Hz hervorruft. Wünschenswert ist auch die Bewahrung al-

ler im Lautstärkeverlauf über das musikalische Geschehen enthaltenen Informationen, also ein konsequent nicht-reduktionistisches Verfahren. Diesen und weiteren hier nicht zu diskutierenden Bedingungen genügt die Fourier-Transformation des Lautstärkeverlaufs, wobei allerdings zwei sehr wichtige Modifizierungen vorgenommen werden müssen: Erstens sind die periodischen Vorgänge mit einem exponentiellen Decay, einer Dämpfung zu versehen, und zweitens ist die Fourier-Transformierte mit jedem Zeitschritt neu, gleichsam stets »aktualisiert« zu erstellen. Auf diese Weise erhält man die sogenannte Zustandsfunktion, die angibt, welcher Oszillator zu welchem Zeitpunkt wie stark aktiviert ist.

Zum Zweck der Veranschaulichung stelle man sich eine Reihe von 145 Pendeln unterschiedlicher Länge vor. Jedes Pendel hat seine Eigenfrequenz in Abhängigkeit von eben dieser Länge. Eine »Riesenfaust« sei in der Lage, all diese Pendel mit einem Schlag gleichzeitig anzustoßen. Diese Faust schlage nun mit einem bestimmten Rhythmus gegen diese Oszillatoren. Je nach Tempo und Rhythmus wird sie manche der Pendel optimal aktivieren (nämlich die, deren Eigenfrequenz der Rhythmus enthält), manche nur mäßig, einige vielleicht auch gar nicht. (Es ist nicht anders als bei einer Kinderschaukel: Den optimalen Schwung erzielt man durch das Anschieben genau zum richtigen Zeitpunkt, bei Wahl eines falschen Zeitpunktes kann die Schaukel zum Stillstand kommen). Auch das erwähnte Decay hat in diesem Bild seinen Platz: Nicht weiter angestoßen wird die Schwingung schwächer und kommt nach einiger Zeit zum Stillstand. Diese Vorstellungen veranschaulichen wesentliche Anteile der mathematischen Vorgänge, die dem Modell zugrunde liegen.

In der musikpsychologischen Literatur ist die Fourier-Transformation gelegentlich anzutreffen, sie wird verwendet beispielsweise bei Gjerdingen (1992, S. 237) oder Todd (1994b, S. 57), diskutiert u.a. bei Rosenthal (1992, S. 73), erschien bislang allerdings nicht in der hier skizzierten speziellen Form.

Von der per Fouriertransformation erstellten Zustandsfunktion gelangt man durch Herausfilterung und Glättung der Aktivierungsmaxima zum sogenannten Tempogramm, welches der Theorie zufolge mit der musikalischen Tempowahrnehmung korrespondiert. Die maximale Abbildungsschärfe eines Tempogramms kann durch geeignete Wahl eines Parameters auf einen bestimmten Bereich der Oszillationsfrequenzen eingestellt, diese Schichten gleichsam fokussiert werden. Die im folgenden diskutierten Abbildungen 3 und 4 enthalten Beispiele für solch unterschiedliche Fokus-

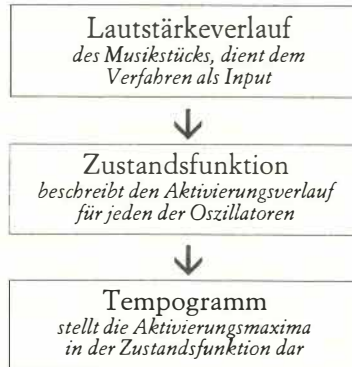


Abb. 2: Schematische Darstellung des Verfahrens.

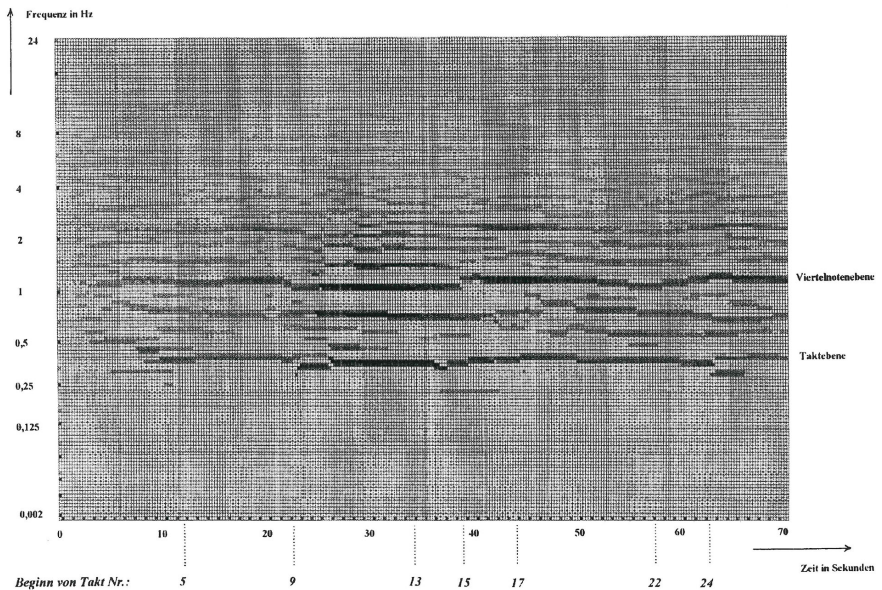


Abb. 3 Tempogramm zur Interpretation von Erik Saties Gymnopédie Nr.1 durch einen Laienpianisten, Fokussierung auf den mittleren Oszillationsbereich. Abgebildet sind die ersten 70 Sekunden, dies entspricht etwa einem Drittel des gesamten Stückes. Jedem der 145 Oszillatoren ist in der Grafik eine horizontale Schicht zugeordnet, die bei starker Aktivierung des betreffenden Oszillators entsprechend geschwärzt ist. Die Zeitachse ist in Sekunden skaliert; zusätzlich sind einige Taktzahlen eingetragen, um einen direkten Bezug auf bestimmte Passagen der Komposition zu ermöglichen. Zwei durchgehende Tempobänder ergeben sich auf Viertel- und Dreiviertelnotebene, wobei letztere in diesem Stück zugleich die Taktebene ist. Die Temposchwankungen des Pianisten bilden sich in den Lageänderungen der Bänder ab.

sierung. Auf diese Weise wird es möglich, besonders interessierende Details auch optisch deutlich herauszuarbeiten.

Abbildung 2 zeigt die einzelnen Schritte des Verfahrens noch einmal in Form eines Flußdiagramms.

Anwendung

Das Verfahren sei demonstriert anhand zweier Einspielungen der Gymnopédie Nr. 1 von Erik Satie:

Betrachten wir zunächst das zur Performance eines Laienpianisten gehörende, auf den mittleren Oszillationsbereich fokussierte Tempogramm in Abbildung 3. Der Ausschnitt umfaßt das in dieser Interpretation ca. 70 Se-

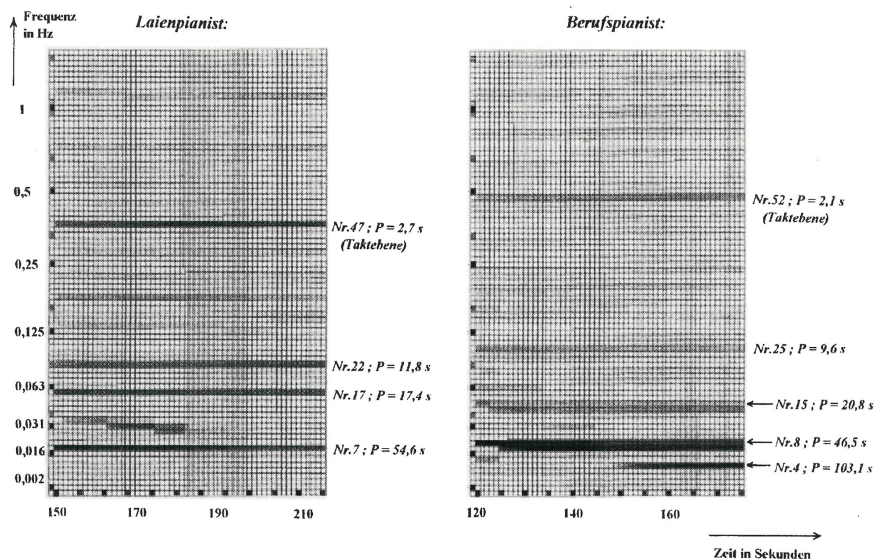


Abbildung 4. Auf den unteren Oszillationsbereich fokussierte Tempogramme zweier Interpretationen im Vergleich: Saties Gymnopédie Nr.1, eingespielt von einem Laien und einem Berufspianisten. Abgebildet ist jeweils die Zeitstrecke, die etwa dem letzten Drittel des Stückes entspricht; in die Darstellung gehen jedoch Daten aus der gesamten Performance ein. Am rechten Rand der Tempogramme sind jeweils die Nummern einiger Oszillatoren und ihre Periode P (die Zeit, die zum Durchlaufen einer vollständigen Oszillation benötigt wird) angeführt. Deutlich an den Schwärzungen zu sehen ist die wesentlich stärkere Aktivierung von niederfrequenten Oszillatoren durch die Interpretation des Berufspianisten, wohingegen der Laie bei seiner Gestaltung wesentlich mehr der Taktebene verhaftet zu sein scheint.

kunden dauernde erste Drittel des Klavierstückes. Das durchgehende dunkle Band knapp oberhalb von 1 Hz zeigt den Tempoverlauf auf der Ebene der Viertelnoten, die mit einem durchschnittlichen Tempo von M.M. = 68 (dies entspricht einer Frequenz von ungefähr 1,13 Hz) gespielt werden. Deutlich zu verfolgen sind die Temposchwankungen auf dieser Ebene, beispielsweise sieht man einen relativ langsam genommenen Abschnitt zwischen den Takten 9 und 14, wohingegen das Maximum im Tempo bei Takt 24 liegt. Anhand der Lage der Aktivierungsmaxima bei ungefähr 1,05 beziehungsweise 1,20 Hz können die Metronomzahlen dieser beiden Tempi näherungsweise ermittelt werden: Wir erhalten die Werte M.M. = 63 und M.M. = 72. Die Ebene der Dreiviertelnoten, dies ist in diesem Stück zugleich die Ganztaktebene, findet sich im Tempogramm als Band unterhalb von 0,5 Hz. Auch hier können Temposchwankungen beobachtet werden, die zumeist parallel zu denen auf Viertelnotenebene verlaufen, innerhalb gewisser Grenzen hat der Spieler jedoch auch die Möglichkeit einer nichtparallelen Gestaltung. Bereits an diesem einfachen Beispiel zeigt sich, daß musikalisches Tempo als ein mehrdimensionales Phänomen aufzufassen ist.

Auch jenseits der Tempogestaltung erlaubt Abbildung 3 nuancierte Aussagen über die Interpretation: Passagen mit starker Schwärzung markieren Abschnitte mit größerer Lautstärke, so in den Takten 9 bis 14 sowie 15 bis 18, wobei sich im ersten dieser Bereiche die stärkere Dynamik auf alle Bereiche erstreckt, im zweiten hingegen auf die Viertelnotenebene beschränkt.

Das Gesamtbild erscheint noch relativ stark gleichsam »verrauscht« durch zahlreiche schwächere Aktivierungsgipfel bei Frequenzen, denen keine in dem Klavierstück vorkommenden Notenwerte entsprechen. Diese Frequenzen sind großenteils als ganzzahlige Vielfache der Oszillationen auf Viertel- und Dreiviertelnotenebene zu verstehen. Ob diesen »Oberschwingungen« ein Korrelat im musikalischen Erleben entspricht, bleibt zu untersuchen. Es wird angestrebt, das Verfahren im Sinne einer komfortablen Lesbarkeit des Tempogramms weiter zu verbessern.

Weitgehend klare Bilder ergeben sich bereits jetzt bei Fokussierung des unteren Bereichs der Oszillationsfrequenzen: Abbildung 4 zeigt die entsprechenden Tempogramme zweier Einspielungen durch den bereits erwähnten Laien und einen Berufspianisten (es wurde die Aufnahme von Frank Glazer, CD Vox 1154862 verwendet). Dargestellt sind hier die Verläufe jeweils im letzten Drittel der Performance. Das Feld oberhalb von 1,5 Hz wurde dabei weggelassen, denn bei Einstellung des Systems auf die

unteren Oszillationsfrequenzen zeigen sich dort keine Aktivierungsmaxima. Zur besseren Orientierung sind die Tempogramme nunmehr am rechten Rand mit einigen Oszillatoren-Nummern und den zugehörigen Perioden versehen. (Die Perioden ergeben sich stets aus dem Kehrwert der Frequenz.)

Man beachte die unterschiedlichen Grundtempi der beiden Interpretationen, die Laienversion dauert ca. 38 Sekunden länger. Für den korrekten Vergleich der beiden Interpretationen bedarf es folglich sozusagen einer »Transposition«: So korrespondiert in der Profi-Version der Oszillator Nr. 52 (mit 0,47 Hz) mit der Taktebene, wohingegen in der langsameren Laienfassung Oszillator Nr. 47 (mit 0,37 Hz) für diese Schicht „zuständig“ ist. Entsprechende Verschiebungen sind auf allen Ebenen zu berücksichtigen.

Der Vergleich der beiden Tempogramme zeigt deutliche Unterschiede: Die Performance des Berufspianisten bringt die wesentlich ausgeprägteren Aktivierungsmaxima an der Untergrenze des Frequenzspektrums hervor. Der Laie erzeugt etwas mehr Anregung im unteren Mittelbereich; dieser Spieler scheint bei seiner Interpretation stärker der Taktebene verhaftet zu sein und um einiges schwächer bei der großräumigen Gestaltung. Beides entspricht dem Höreindruck.

Die meisten der angeregten Oszillationen können bestimmten formalen Einheiten der *Gymnopédie* zugeordnet werden: Die Periode des Oszillators Nr. 15 beispielsweise entspricht mit ca. 20,8 Sekunden ungefähr der Dauer von neun Takten in der Interpretation des Berufspianisten; tatsächlich enthält das Stück zahlreiche acht-, neun- oder zehntaktige Einheiten. Die ebenso häufigen Vier- und Fünftaktgruppen der Komposition spiegeln sich in einer Anregung von Oszillator Nr. 25 mit einer Periode von 9,6 Sekunden. Beim Tempogramm des Berufspianisten – und bemerkenswerterweise nur dort – kann auch die Gliederung der Komposition in zwei weitgehend identische Hälften mit der Aktivierung eines Oszillators (Nr. 4) in Verbindung gebracht werden. Dessen Periode von 103 Sekunden liegt zwar um etwa 15 Sekunden über der halben Gesamtlänge des Stückes, jedoch ist zu berücksichtigen, daß das System im unteren Frequenzbereich infolge des großen Abstandes der Oszillatoren zueinander (hier unterscheiden sich zwei Nachbarn in ihren Frequenzen um erheblich mehr als 4%) mit einiger Ungenauigkeit arbeiten muß. Die direkten Nachbarn des 103-Sekunden-Oszillators haben Perioden von 80 bzw. 140 Sekunden, ein Unterschied von 15 Sekunden ist also in diesem Bereich ein geringerer als der zwischen zwei benachbarten Einheiten.

Die Ergebnisse der hier vorgestellten neuen Methode zur Performanceanalyse erfüllen die weiter oben aufgestellten Forderungen: Erstens wird die Mehrdimensionalität der Tempogestaltung erfaßt und in sinnfälliger Weise dargestellt. Zweitens eröffnet sich infolge des nicht-reduktionistischen Ansatzes – es werden vollständige Stücke analysiert – und der Einbeziehung sehr niederfrequenter Oszillatoren die Möglichkeit, etwas über großräumige Zusammenhänge auszusagen. Durch den Vergleich einer Laien- mit einer Expertenperformance konnte die Vermutung erhärtet werden, daß sich gerade in diesem Oszillationsbereich Qualitätsmerkmale einer Performance zeigen. Wir bewerten diese ersten Ergebnisse als Ermutigung zu einer Weiterentwicklung und einem Ausbau des Verfahrens.

Perspektiven

1. Zunächst gilt es, die grafische Darstellungsform des Aktivierungsmusters durch eine Feinabstimmung entsprechender Optionen so zu regulieren, daß der »Störeffekt des Rauschens«, wie er in den oberen Frequenzbändern von Abbildung 3 zu sehen ist, minimiert wird.

2. Mit dem Erreichen der optimalen Darstellungsform wird es dann möglich sein, den Verlauf einzelner Bänder von Aktivierungsgipfeln in ein quantifizierbares Maß für die vom Spieler beispielsweise auf der Viertelnoten-Ebene verwendete Agogik zu überführen. Durch dieses Maß – wir wollen es vorerst Agogizität nennen – wäre es möglich, einen Interpretationsvergleich durchzuführen, der die verschiedenen Interpretationen durch konkrete Zahlenwerte für die Agogik auf einer bestimmten Notenwertebene beschreibt.

3. Für den Interpretationsvergleich erschließen sich völlig neue Möglichkeiten, denn mit dem grafischen Ausdruck in Form des sogenannten Tempogramms kann nun – neben dem Vorhandensein niederfrequenter Oszillationen – ein weiteres Qualitätskriterium für Interpretationen gebildet werden: Ausgehend von der Annahme, daß eine gute Performance sowohl durch Differenzierung (z.B. durch Agogik auf Notenwertebene) als auch durch Integration (d.h. durch eine Perspektive, die den strukturellen Gesamtverlauf des Stückes im Blick hat) gekennzeichnet ist, läßt sich die Hypothese formulieren, daß Expertenpianisten einen höheren Komplexitätsgrad der Zeitgestaltung erreichen als Laien. Durch weitere Analysen werden wir diese Hypothese prüfen.

4. Eine Erweiterung des Verfahrens wird zukünftig auch andere Parameter einer Performance analysierbar machen. Geplant ist derzeit die Analyse auch der rhythmischen Qualität einer Interpretation und die ihrer dynamischen Gestaltung. Langfristig wird angestrebt, auch für die rhythmische und dynamische Ausführung eines Werkes quantifizierbare Beschreibungsmaße zu entwickeln, denen wir vorerst die Bezeichnungen Rhythmisizität und Dynamizität geben wollen.

5. Die Verwendung eines MIDI-Klaviers wird es ermöglichen, beide Hände eines Pianisten getrennt zu erfassen, und so die Komplexität mehrerer nicht-synchronisierter, jedoch simultan verlaufender Rubatoverläufe zu analysieren.

6. Die äußerst sinnfällige Darstellung der im Spiel enthaltenen Zeitinformationen durch das Tempogramm kann pädagogisch genutzt werden, indem dem Spieler durch die entsprechende Software zukünftig ein Feedbackinstrument zur Verfügung gestellt wird, das es ihm ermöglicht, mittels eines PCs beispielsweise die Umsetzung seiner Phrasierungsintention objektiv zu überprüfen. Sehr groß gedachte Phrasenbögen z.B. müßten – bei gelungener Realisierung – im Tempogramm als niederfrequente Oszillationen sichtbar werden.

7. Da das Berechnungsverfahren unempfindlich gegenüber akustischen Störfaktoren wie Knistern oder Rauschen ist, können selbst historische Aufnahmen mit schlechter Klangqualität als Input verwendet werden. Hiermit wird es möglich sein, die Performanceanalyse auf einen historischen Abschnitt auszudehnen, der bisher weitgehend unerforscht blieb.

8. Die Tätigkeit des Spielers erscheint vor dem Hintergrund der auf den unteren Frequenzbereich fokussierten Tempogramme in einem neuen Licht: Der Sinn einer Performance erschöpft sich nicht in der »Anreicherung« des Notentextes mit sogenannten "expressive deviations", sondern die bisherigen Tempogramme lassen vermuten, daß eine gute Performance auch formbildende Aspekte beinhaltet (vgl. die Analysen zu Abbildung 4). Es zeigte sich, daß den meisten der niederfrequenten Oszillationen eindeutig Formabschnitte zugeordnet werden können. Dies ist eine Bestätigung der Ausdruck-Struktur-Theorie von Clarke (z.B. 1985), die besagt, daß ein sinnvoller Ausdruck in einer Performance nur durch einen Bezug auf die kompositorische Struktur entstehen kann. Eine gute Performance hätte somit eine sinnstiftende Funktion.

Bemerkenswerterweise arbeitet der TOS-Algorithmus so, daß niederfrequente Oszillationen nur dann entstehen, wenn die Details einer Perfor-

mance über weite Strecken aufeinander bezogen sind. Im Extremfall bedeutet dies: Bereits zu Beginn eines Stückes muß der Spieler wissen, welche Gestaltung er bis zum Ende vornehmen wird, und auch am Ende ist ihm noch die anfängliche Konzeption seines Vortrags präsent. Eine gute Performance wäre demzufolge ein in hohem Maß planvolles Tun – mag dies auch unbewußt geschehen.

9. Aus den bisherigen Analysen ergibt sich, daß die Anzahl der Oszillatoren für sehr niedere Frequenzen in der weiteren Feinabstimmung des Modells erhöht werden muß, um damit die Darstellungsgenauigkeit zeitlicher Verläufe zu verbessern. Eine Genauigkeit wie etwa bei den Frequenzen um 1 Hz ist jedoch nicht anzustreben, da die menschliche Wahrnehmung bei Zeitstrecken in der Größenordnung von Minuten ebenfalls nicht mit der gleichen Präzision wie im Sekundenbereich arbeitet.

10. Mit der TOS wird auch die Frage: »Wie lang ist eine Phrase?« beantwortbar. Diese Frage hat im musiktheoretischen Schrifttum eine mittlerweile 100jährige Tradition, und die Antworten fielen – je nach Autor – unterschiedlich aus. Man kann im wesentlichen drei Ansätze benennen, mit denen eine Antwort versucht wurde: die von Riemann (1895), Schenker (1925) und Cooper & Meyer (1960). Ihre Grundgedanken seien im folgenden skizziert:

In Riemanns Denken hat eine Phrase einen eher geringen Umfang. Ausgehend vom einzelnen Motiv wird dieses innerhalb einer Phrase fortgeführt. Hierbei erreicht die Phrase eine Länge von vier oder etwas mehr Takten. Das Begrenzungskriterium für eine Phrase ist die Ausprägung eines Tonikaraumes durch einen Halb- oder Ganzschluß (1895, S. 69). Die nächstgrößere Bezugseinheit ist dann die Periode. In ihr werden Phrasen zusammengefaßt (1895, S. 86). Läßt man den Phrasenbegriff auch für die Periode gelten, so ergibt sich eine maximale Ausdehnung von acht Takten.

In Schenkers analytischem Ansatz gibt es keine vorgegebene Begrenzung musikalischer Einheiten auf eine festgelegte Größe. Ganz im Gegenteil erstrecken sich Ursatz und Urlinie (als ausgefüllte Oberstimme des Ursatzes) über das ganze Stück. Diesen unteilbaren, übergeordneten harmonisch-melodischen Zug arbeitet Schenker an einer Vielzahl von Beispielen heraus: In seinem Buch »Das Meisterwerk in der Musik« zeigt er auf, wie sich die Urlinie über das gesamte Präludium aus Bachs E-Dur-Partita für Violine Solo erstreckt (1925, S. 77ff.). Die Urlinie kann sich jedoch auch über den ganzen Satz einer Sinfonie erstrecken, was Schenker an Beethovens Eroica nachweist (1925, S. 25ff.). Diese Phrasenzüge sind von großer Relevanz für

den musikalischen Vortrag, denn »dem Vortragenden ist die Urlinie zunächst eine Richtung, etwa das, was dem Bergsteiger eine Wegkarte ist (...)« (Schenker 1925, S. 195). In Schenkers Vortragslehre ist deshalb ein Vortrag, der nicht gleichzeitig das Ganze (die Urlinie) und das Detail (die Diminution) im Blick hat, verfehlt (Schenker 1925, S. 196). Schenker, der selber ein ausgezeichneter Pianist war, hat stets auf die Einheit von Analyse und Vortrag hingewiesen, um »blindes« Spiel zu vermeiden. Aufschlußreich sind hier seine Eintragungen in Beethovens Klaviersonaten, durch die deutlich wird, wie nach Schenkers Vorstellung der Vortrag im Spannungsfeld von Großform und Detail oder – mit anderen Worten – von Integration und Differenzierung zu gestalten sei (vgl. Rothstein 1984). Folgt man der These von Gurlitt (1954, S. 653), die Kompositionsgeschichte der europäischen Musik seit dem Mittelalter sei geprägt durch die drei zeitgestaltenden Prinzipien mensura, tactus und tempo, dann ist folgerichtig anzunehmen, daß der Vortrag als Realisation der einkomponierten Zeitgestaltung dieser entsprechen muß.

Aus anderer Richtung nähern sich Cooper & Meyer (1960) der Frage nach der möglichen Länge einer Phrase. Sie gehen von einer eher traditionellen Versfußtheorie des musikalischen Rhythmus aus und nehmen eine Vielzahl möglicher rhythmic levels an, die innerhalb eines Stückes existent sein können. Wir finden auch hier die Idee einer multiplen Zeitarchitektur ausgeprägt. Die rhythmischen Ebenen werden dann stets durch betonte bzw. unbetonte Versfüße ausgefüllt. Die ausgedehnteste rhythmische Ebene umfaßt auch hier maximal die gesamte Länge eines Stückes, und in der Analyse von Schönbergs Klavierstück op. 19, 4 zeigen die Autoren auf, wie sich der dem ganzen Stück zugrundeliegende Versfuß auf das ca. 30 Sekunden dauernde Stück verteilt (1960, S. 175ff.). Hierdurch ergäbe sich in unserem Modell eine Oszillation von 0,033 Hz. Im ersten Satz von Beethovens 8. Sinfonie analysieren Cooper & Meyer (1960, S. 203ff.) auf dem rhythmischen »level 5b« dann einen »Super«-Versfuß, der sich über das ca. neun-einhalbminütige Stück erstreckt. Dies entspräche einer Oszillation von 0,0017 Hz. Die ersten 103 Takte der Exposition stellen eine (unbetonte) Senkung dar, die Wiederholung der Exposition ist dann die entsprechende (betonte) Hebung, und die Takte 104–373 (Durchführung und Reprise) bilden eine erneute Senkung.

Die TOS erweist sich an diesem Punkt als ein theoretisches Gebäude, das die bisherigen, sehr verschiedenen Erklärungsversuche zur Phrasenlänge widerspruchsfrei in sich aufzunehmen vermag: Jeder der zitierten Autoren

beschreibt dieser Theorie zufolge Teilaspekte eines durch die TOS vollständig erfaßbaren Ganzen. Musikalische Bausteine jeder Länge – bestehen sie aus wenigen Noten oder ganzen Satzteilen – korrespondieren mit Oszillationen, deren Schwingungsdauer eben dieser Länge entspricht. Die gesamte formale Architektur eines Musikstückes mit allen Einheiten und Subeinheiten auf den verschiedenen zeitlichen Ebenen spiegelt sich in den Aktivierungsmustern der Zustandsfunktion und ist dort der Beobachtung zugänglich.

Auch von den erwähnten Autoren wird die Notwendigkeit, musikalische Zeitgestaltung als mehrdimensionales Phänomen aufzufassen, mehr oder weniger deutlich gesehen; die TOS stellt jedoch insoweit eine Fortentwicklung dar, als da hier erstmals die vollständige Erfassung aller Schichten – vom kleinsten Notenwert bis hin zur Dauernebene des gesamten Stückes – möglich wird.

Bereits durch die kompositorischen Struktur eines Werkes ist in der Regel diese hochkomplexe Architektur gegeben; man mag nun vermuten – und unsere ersten Untersuchungen weisen in diese Richtung –, daß eine gute Performance auf diese kompositorische Vieldimensionalität mit einer »kongenialen« interpretatorischen Vieldimensionalität antwortet.

Zusammenfassung

Neue Möglichkeiten der Performance-Analyse auf Grundlage einer Theorie oszillierender Systeme (TOS) werden vorgestellt. Dieser Theorie zufolge löst der Lautstärkeverlauf eines Musikstückes in der menschlichen Wahrnehmung eine Reihe von Oszillationen zwischen 0,002 und 24 Hz aus, deren jeweilige Stärke in Abhängigkeit vom konkreten Input mittels einer modifizierten Fourier-Transformation berechnet werden kann. Die graphische Darstellung solcher Aktivierungsstärken, das sogenannte Tempogramm, erlaubt differenzierte Aussagen über Eigenheiten und Qualität einer Performance. Demonstriert wird das Verfahren anhand zweier Einspielungen der *Gymnopédie* Nr. 1 von Erik Satie. Die Tempogramme dieser beiden Interpretationen lassen zum einen die feinen Schwankungen in der Tempogestaltung erkennen, sie zeigen zum anderen, daß dem Berufspianisten im Gegensatz zum Laien die Aktivierung wesentlich stärkerer Oszillationen im niederfrequenten Bereich gelingt, was einer Hypothese der TOS zufolge ein Qualitätskriterium für die Interpretation darstellt. Die Untersuchungen er-

geben stets die simultane Anregung von mehreren Oszillator-Schichten und stützen somit die Ansicht, daß musikalisches Tempo grundsätzlich als mehrdimensionales Phänomen aufzufassen ist. Forschungshistorische Verknüpfungen des Projektes werden dargestellt sowie Perspektiven für die weitere Arbeit mit diesem Ansatz aufgezeigt.

Summary

This paper presents a new method of performance analysis based on a “Theory of Oscillating Systems” (TOS). The theory assumes that the dynamic course of a performance triggers a number of oscillations – situated between 0.002 and 24 Hz – (perceived by non-adaptive oscillators) in human perception. The oscillations’ specific activation depends on the input source and can be evaluated by a modified Fourier analysis. These states of activation are shown in a diagram – the so called Tempogram – which enables us to draw significant conclusions about the peculiarities and quality of a performance. A demonstration of the method is given, using two recordings of Erik Satie’s *Gymnopédie No. 1* played by a professional and a non-professional pianist. Tempograms of both interpretations show, on the one hand, fine deviations in tempo design and, on the other, that the professional pianist’s playing is characterized by more intensive activation of the oscillations situated in the low-frequency area. This result confirms a hypothesis of the “Theory of Oscillating Systems” – namely that low-frequency oscillations provide a criterion for measuring the quality of an interpretation. The analyses always show the simultaneous stimulation of various layers of oscillators; this confirms our assumption that musical tempo has to be seen basically as a multidimensional phenomenon. Perspectives for future research and connections to other methods of performance research and musical phrasing are discussed.

Anmerkungen

Unser Dank gilt folgenden Personen: Dr. Bernhard Feiten (Technische Universität Berlin) für die Einrichtung und Einweisung in das Audioanalyseprogramm CARL, Prof. Dr. Klaus-Ernst Behne (Musikhochschule Hannover) für seine wertvollen Verbesserungsvorschläge zu einer frühen Manuskriptfassung, Prof. Dr. Horst-Peter Hesse (Salzburg) für die Überlassung seines Manuskripts zur Tempowahrnehmung und Andreas Bernnat für die Bereitschaft, die Laienversionen der analysierten Stücke einzuspielen.

Literatur

- Battel, G. U., Bresin, R., De Poli, G. & Vidolin, A. (1993) – *Automatic performance of musical scores by means of neural networks: Evaluation with listening tests*. Atti del X Colloquio di Informatica Musicale, Milano, 2.-4. December.
- Clarke, E. F. (1985) – *Structure and expression in rhythmic performance*. In: P. Howell & R. West (Ed.), *Musical structure and cognition*. London: Academic Press.
- Cooper, G. W. & Meyer, L. (1960) – *The rhythmic structure of music*. Chicago: Academic Press.
- Desain, P. & Honing, H. (1994) – *Can music cognition benefit from computer music? – From foot tap-per systems to beat induction models*. Proceedings of the International Conference for Music Perception and Cognition, Liege, 23. 27. July, 397–398.
- Drake, C., Botte, M.C. & Baruch, C. (1992) – *Tempo sensitivity: Regularity detectors?* Proceedings of the Fourth Workshop on Rhythm Perception & Production, Bourges, France, June, 1722.
- Epstein, D. (1985) – *Tempo relations: A crosscultural study*. Music-Theory Spectrum, 7, 34–71.
- Gabrielsson, A. (1987) – *Once again. The theme from Mozart's piano sonata in A major (K. 331)*. A comparison of five performances. In A. Gabrielsson (Ed.) – *Action and perception in rhythm and meter* (Publications issued by the Royal Swedish Academy of Music, No. 55). Stockholm.
- Gjerdingen, R. O. (1992) – *Revisiting Meyer's "Grammatical simplicity and relational richness"*. In: M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication*. Washington: American Psychological Association.
- Gjerdingen, R. O. (1993) – *"Smooth" rhythms as probes of entrainment*. Music Perception, 10, (4), 503–508.
- Gottschewski, H. (1993) – *Theorie und Analyse der musikalischen Zeitgestaltung*. Neue Wege der Interpretationsforschung, gezeigt an Welte-Mignon-Aufnahmen aus dem Jahre 1905 (Dissertation, Universität Freiburg).
- Gurlitt, W. (1954) – *Form in der Musik als Zeitgestaltung*. Akademie der Wissenschaften und der Literatur Mainz: Abhandlungen der Geistes- und Sozialwissenschaftlichen Klasse, 13, 651–677.
- Hartmann, A. (1932) – *Untersuchungen über metrisches Verhalten in musikalischen Interpretationsvarianten*. Archiv für die gesamte Psychologie, 84, 103–192.
- Jones, M. R. & Yee, W. (1993) – *Attending to auditory events: The role of temporal organization*. In: S. McAdams & E. Bigand (Eds.), *Thinking in sound. The cognitive psychology of human audition*. Oxford: Clarendon Press.
- Kopiez, R. (1995) – *Aspekte der Performanceforschung*. In: H. de la Motte-Haber, *Handbuch der Musikpsychologie*, (2. Auflage). Laaber.
- Langner, J. (i.V.) – *Musikalischer Rhythmus und Oszillation*. Eine theoretische und empirische Erkundung (Dissertation, Musikhochschule Hannover).
- Large, E. W. (1994) – *Dynamic representation of musical structure* (Dissertation, Ohio State University).
- Large, E. W. & Kolen, J. F. (1994) – *Resonance and the perception of musical meter*. Connection Science, 6, (2 & 3), 177–208.
- Mazzola, G. (1993) – *Geometry and logic of musical performance* (Swiss National Science Foundation, Grant Nr. 21 33651.92). Report. Zürich: Author.
- Mazzola, G. & Zahorka, O. (1994) – *Tempo curves revisited: Hierarchies of performance fields*. Computer Music Journal, 18, (1), 40–52.
- Parncutt, R. (1994) – *A perceptual model of pulse salience and metrical accent in musical rhythms*. Music Perception, 11, (4), 409–464.
- Povel, D.J. & Essens, P. (1985) – *Perception of temporal patterns*. Music Perception, 2, (4), 411–440.
- Repp, B. (1990) – *Patterns of expressive timing in performances of a Beethoven minuet by 19 famous pianists*. Journal of the Acoustical Society of America, 88, (2), 622–641.
- Repp, B. (1992) – *Diversity and communality in music performance. An analysis of timing microstructure in Schumann's »Träumerei«*. Journal of the Acoustical Society of America, 92, (5), 2546–2568.

- Riemann, H. (1895) – *Die musikalische Phrasierung* (Präludien und Studien: Bd. I). Leipzig.
- Rosenthal, D. (1992) – *Emulation of human rhythm perception*. Computer Music Journal, 16, (1), 6476.
- Rothstein, W. (1984) – *Heinrich Schenker as an interpreter of Beethoven's Piano Sonatas*. 19th Century Music, 8, (1), 327.
- Schenker, H. (1925) – *Fortsetzung der Urlinie-Betrachtungen*. (Das Meisterwerk in der Musik: Bd. I). München.
- Seashore, C. E. (1938). *Psychology of music*. New York: Dover Publications. Reprint 1967.
- Seashore, C. E. (Ed.). (1937) – *Objective analysis of music performance* (University of Iowa, Studies in the psychology of music: Vol. 4). Iowa: The University Press.
- Shaffer, L. H. (1980) – *Analysing piano performance*. A study of concert pianists. In: G.E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North Holland Publishing Company.
- Shaffer, L. H. (1981) – *Performances of Chopin, Bach, and Bartok*. *Studies in motor programming*. Cognitive Psychology, 13, 326–376.
- Sundberg, J. (1988) – *Computer synthesis of music performance*. In: J. A. Sloboda (Ed.), *Generative processes in Music. The Psychology of Performance, Improvisation, and Composition*. Oxford: Clarendon Press.
- Sundberg, J. (1991) – *Music performance research*. An overview. In: J. Sundberg, L. Nord & R. Carlson (Eds.), *Music language, speech and brain*. London: Macmillan.
- Todd, N. P. M. (1994a) – *Metre, grouping and the uncertainty principle: A unified theory of rhythm perception*. Proceedings of the International Conference for Music Perception and Cognition, Liège, 23.–27. Juli, 395–396.
- Todd, N. P. M. (1994b) – *The auditory "primal sketch": A multiscale model of rhythmic grouping*. Journal of New Music Research, 23, 25–70.