Hatte Lipps doch recht?
Tonverwandtschaft und Tonverschmelzung im Lichte der heutigen Gehörphysiologie

Horst-Peter Hesse

Verschmelzung und Gestaltqualitäten

Eine Gruppe von Tönen wird unter bestimmten Bedingungen als zusammengehörige Einheit aufgefaßt. Besonders deutlich wird dies bei gleichzeitig erklingenden Tönen, die - je nach dem Verhältnis der Schwingungszahlen - in unterschiedlichem Grade zu wohlklingenden Einheiten verschmelzen können. Das gilt in weiterem Sinne nicht nur für simultan erklingende Töne, für Intervalle und Akkorde, sondern für Melodien gleichermaßen; man nimmt sie als prägnante Gebilde wahr, als geschlossene und zugleich strukturierte Einheiten, die durch eine spezifische Klangqualität ausgezeichnet sind. Solche Tonverbindungen haben Eigenschaften, die an den einzelnen Tönen nicht in Erscheinung treten. Die neuen Eigenschaften entstehen aber nicht durch Mischung der beteiligten Toneigenschaften, denn die Einzeltöne behalten ihre Identität; diese gehen nicht - wie gemischte Farben - in dem Verschmelzungsprodukt auf sondern können bei aufmerksamem Hinhören als Elemente des Ganzen herausgehört werden. Darüber hinaus aber ist durch die Zusammenfügung der Töne etwas Neues entstanden, eine Tongestalt, die durch eigene Merkmale charakterisiert ist.

Der Philosoph Christian Freiherr von Ehrenfels lenkte den Blick durch seine programmatische Schrift auf die seit der Antike bekannte These: "Das Ganze ist mehr als die Summe seiner Teile" [1]. Intervalle, Akkorde, Motive und Melodien lassen sich innerhalb des Tonsystems transponieren, wobei jeder Ton durch einen anderen ersetzt wird, ohne daß das Ganze seinen spezifischen Charakter verliert. Für diejenigen Eigenschaften, die trotz Transposition unverändert erhalten bleiben, prägte Ehrenfels den Begriff *Gestaltqualitäten.* Natürlich war die Existenz von Gestalten seit jeher bekannt, aber die Frage, wie es im Prozeß der Wahrnehmung zur Entstehung von Gestalten kommt, wurde nach dem Anstoß, den Ehrenfels gegeben hatte, neu durchdacht. Während man bis dahin davon ausging, daß Zusammenhänge zwischen Einzelempfindungen durch bewußte psychische Aktivität erzeugt würden, hatte bereits Ernst Mach den Gedanken geäußert, daß wir Tongestalten, wie z.B. Melodien, unabhängig von jeglicher intellektueller Verarbeitung unmittelbar als solche wahrnehmen könnten [2]. Als revolutionärer Kerngedanke der neuen Gestalttheorie reifte allmählich die These, daß Gestalten nicht sekundäre Produkte sind, die das Subjekt mehr oder weniger willentlich aus den einzelnen Wahrnehmungselementen zusammensetzt, sondern daß in sich geschlossene Ganzheiten im Erlebnis primär gegeben sind. Dies wurde als "spontane gestalthafte Organisation" bezeichnet.

Natürlich haben sich an dieser Behauptung lebhafte wissenschaftliche Kontroversen entzündet, obgleich die Alltagserfahrung diese Aussage stützt: Wir sehen um uns herum Dinge und Ereignisse und nicht etwa ein Raster von verschiedenen Farbtönen in unterschiedlicher Helligkeit. Lausche ich dem Spiel auf einem Musikinstrument, so höre ich kein Spektrum von Sinustönen sondern eine Melodie in einer bestimmten Klangfarbe. Die Klangfarbe erkennt man, auch wenn man die konstituierenden Teiltöne nicht identifizieren kann. Man kann Rhythmen erfassen und im Gedächtnis behalten, ohne die einzelnen Notenwerte zu wissen, Melodien entsprechend, ohne über Einzeltöne bzw. Intervalle Rechenschaft ablegen zu können. Aus den Gestalten kann man nachträglich - bei analytischer Einstellung - Elemente ausgliedern, auf Grund derer rationale Aussagen über das Gehörte erleichtert werden oder das Gehörte aufgeschrieben werden kann.

Daß wir dies so erleben, wurde zwar von niemandem bezweifelt, wohl aber gab es unterschiedliche Antworten auf die Fragen, ob ein physiologischer oder ein psychischer Ordnungsprozeß der bewußten Wahrnehmung vorangegangen sei und ob bzw. in welchem Umfange die gestaltbildenden Tendenzen angeboren oder von Traditionen und Lernprozessen abhängig sind oder gar durch Lenkung der Aufmerksamkeit willentlich beeinflußt werden können. Ist beispielsweise das Grundtonempfinden bei Melodien oder Zusammenklängen naturgegeben

oder durch Gewöhnung entstanden, liegen die Kriterien für die Qualität der Intervalle in der klingenden Materie oder in Struktur und Arbeitsweise des Nervensystems, oder spielen hier Konventionen eine entscheidende Rolle? Eine Antwort auf die Frage, ob die beobachteten Wechselwirkungen bereits auf der physischen oder erst auf der psychischen Ebene stattfinden, ist insofern von großer Bedeutung, als Wechselwirkungen, die schon im physischen Bereich stattfinden, der Willensentscheidung entzogen sind und daher als anthropologische Konstanten gelten können.

Carl Stumpf hat diese Fragen in seiner *Tonpsychologie* eingehend diskutiert [3]. Es gibt keine gründlichere Abwägung der Argumente für bzw. gegen die verschiedenen Erklärungen der Tonverschmelzung. Alle psychologischen Erklärungsversuche, so stellte er fest, müssen mißlingen, denn die Verschmelzung ist eine Empfindung, die von Erfahrung und Übung im individuellen Leben unabhängig ist. Ebensowenig sind peripherienahe, physiologische Erklärungen möglich, weil einerseits zusammengesetzte Schwingungen im Innenohr in Komponenten aufgelöst werden und weil andererseits eine Verschmelzung auch dann stattfindet, wenn die Töne getrennt auf beide Ohren geleitet wurden, so daß also ein zentraler Prozeß erforderlich ist.

> *"Es müssen also den Unterschieden der Verschmelzungsgrade gewisse Unterschiede der letzten Vorgänge im Hörzentrum als physisches Korrelat oder als Ursache (jenachdem man monistisch oder dualistisch denkt) entsprechen. Welcher Art aber diese Unterschiede sind, darüber wissen wir schon darum nichts, weil wir über die Beschaffenheit der letzten Prozesse überhaupt nichts wissen."* [4]

Sowohl die neuralen Zentren des Hirnstamms als auch das Großhirn sind unvorstellbar komplizierte dreidimensionale Netzwerke, in denen die einzelnen Nervenzellen mit Tausenden teils nah benachbarten, teils weit entfernten anderen Nervenzellen verknüpft sind. Daher sind die Vorgänge in den vom Innenohr zur Großhirnrinde führenden Nervenbahnen derart kompliziert und schwierig zu untersuchen, daß zwischen gesicherten Fakten noch immer Lücken bestehen, die durch Hypothesen überbrückt werden müssen. Und über deren Gültigkeit ist man sich unter den Forschern naturgemäß nicht einig.

Zur Frequenzanalyse im Innenohr

Die hypothetischen Erklärungsmodelle lassen sich auf zwei unterschiedliche Prinzipien zurückführen, in denen die Vertreter zweier divergierender Forschungsrichtungen die Lösung von grundlegenden Problemen des Hörens vermuten. Das erste Prinzip liegt der sogenannten *Ortstheorie* des Hörens zugrunde, die - ausgehend von der Helmholtzschen Resonanztheorie [5] - in unserem Jahrhundert vor allem durch die bahnbrechenden Forschungen des Nobelpreisträgers Georg von Békésy weiterentwickelt wurde [6]. Die meisten Wissenschaftler haben sich heute für die *Ortstheorie* entschieden, da die gehörphysiologische Forschung die von Helmholtz begründete Theorie zwar inzwischen wesentlich modifiziert, aber ihre Grundkonzeption für den peripheren Anteil des Hörsystems eindeutig bestätigt hat. Wir wissen heute, daß im menschlichen Innenohr eine Frequenzanalyse stattfindet, als deren Ergebnis die Schallwellen - je nach ihrer Frequenz - an unterschiedlichen Orten der Cochlea Nervenimpulse auslösen, und daß komplexe Schallwellen durch eine ihren Frequenzkomponenten entsprechende, örtlich aufgefächerte Erregungsverteilung repräsentiert werden. Diese im Innenohr beginnende soge-nannte Tonotopie läßt sich im gesamten auditorischen System bis in die Hirnrinde verfolgen. Die Verarbeitung der Nervenimpulse in den höheren Zentren der Hörbahn entzieht sich jedoch dem Zugriff der Forschung noch immer weitgehend. Daher versuchte man, so viele Hörphänomene wie möglich durch peripherienahe Prozesse zu erklären.

Viele Beobachtungen können allerdings bis heute nicht in befriedigendem Umfange erklärt werden. Das gilt insbesondere für eine Reihe von Wechselwirkungen, die bei gleichzeitig erklingenden Tönen auftreten. Die ungeklärten bzw. bisher kontrovers beantworteten Fragen beginnen jedoch schon bei der Wahrnehmung des Einzeltones. Die Tonhöhenwahrnehmung ist nicht etwa bei sinusförmigen Schallwellen besonders präzise, im Gegenteil: Die Genauigkeit, mit der die Tonhöhe bestimmt werden kann, steigt mit der Anzahl der harmonischen Komponenten einer Schallwelle (bis zu einem Grenzwert), obgleich die Beurteilung durch die Obertöne doch

eigentlich erschwert werden müßte, da diese eine Reihe sich kontinuierlich verengender Intervalle über dem Grundton bilden. Eine einheitliche Tonhöhe kann sogar dann gehört werden, wenn die Grundfrequenz im Schwingungsvorgang objektiv gar nicht vorhanden ist. Diese Tonhöhenempfindung, der keine Grundfrequenz im Schallspektrum gegenübersteht, wurde von ihrem Entdecker, dem Holländer Schouten, *Residuum* genannt [7]. Er fand, daß die höheren Harmonischen nicht einzeln aufgefaßt werden können, sondern daß sie kollektiv wirksam werden mit einer Tonhöhe, die der Periodendauer der von ihnen gemeinsam gebildeten Wellenform entspricht. Diese ist gleich derjenigen der Grundschwingung. Das würde bedeuten, daß die Residualtonwahrnehmung auf Reizung von Orten in der Cochlea beruht, deren Erregung durch sinusförmigen Schall sonst die Wahrnehmung hoher Töne bewirkt.

Zur Erklärung dieses Phänomens lieferte Ernst Terhardt ein Modell auf der Basis der Ortstheorie, das wegen seiner Evidenz sehr schnell Eingang in das Schrifttum fand [8]: Wenn ein Ton erstens nicht durch einen einzigen Ort in der Cochlea repräsentiert wird sondern - wegen seiner Zusammensetzung aus Partialtönen - durch die von verschiedenen Orten ausgehenden Nervenimpulse, und wenn diese Orte zweitens - wegen des regelmäßigen Aufbaus der Partialtonreihe - in festen Distanzverhältnissen zueinander stehen, dann könnte die Tonhöhenerkennung als ein erlernter Mustererkennungsprozeß erklärt werden, der auch bei unvollständigem Muster (z.B. bei fehlendem Grundton) aus der Konstellation der übrigen Teiltöne die Tonhöhe erkennt, so wie auch unvollständige optische Zeichen erkannt werden. Terhardt nennt die Tonhöhenempfindung in diesem Falle *Virtuelle Tonhöhe*.

Auch die musikalisch fundamentale Unterscheidung von Konsonanz und Dissonanz hat Terhardt sehr gründlich beleuchtet [9]. Sein Konzept schließt eng an die Helmholtzsche Theorie an, wo es heißt: "Wenn zwei musikalische Klänge nebeneinander erklingen, ergeben sich im allgemeinen Störungen ihres Zusammenklingens durch die Schwebungen, welche ihre Partialtöne miteinander hervorbringen, so daß [...] der Zusammenklang rauh wird." [10] Es mußte aber ein weiterer Faktor berücksichtigt werden. Schon Stumpf hatte beobachtet, daß die als optimal empfundene Intonation eines Intervalles davon abhängt, ob es simultan oder sukzessiv dargeboten wird. Bei simultaner Darbietung mehrerer Schallschwingungen beeinflussen die einzelnen Komponenten sich offenbar gegenseitig in der Weise, daß die frequenzspezifischen Erregungsmaxima sich gegenüber ihrer Lage bei Einzeldarbietung auseinanderschieben.

Tritt also zu einem Grundton dessen Oktave hinzu, so spreizt sich der Abstand zwischen den beiden Erregungsmaxima, der - nach der Ortstheorie - als Tondistanz, als Intervall bewußt wird. Durch die Erfahrung mit stimmhaften Sprachlauten, deren erste sechs harmonische Komponenten einen über zweieinhalb Oktaven ausgebreiteten Durdreiklang bilden, erlange das Gehör - so Terhardts Hypothese - eine Kenntnis von Tonhöhendistanzen, die sich als Normwerte einprägen. Erklingen nun beispielsweise Grundton und Oktave nicht gleichzeitig sondern nacheinander, so verringere sich der Abstand der Erregungsmaxima gegenüber der für die Oktave vorgespeicherten Norm; das Gehör verlange daher eine Erhöhung des Oktavtones. Aus dieser Sicht könne die Sonderstellung, die die einfachen Zahlenverhältnisse seit Jahrtausenden genießen, nicht mehr aufrechterhalten werden.

Terhardt modifizierte die traditionelle Theorie nun dahingehend, daß er für die Bestimmung der Intervallgrößen nicht die Verhältnisse kleiner ganzer Zahlen gelten ließ, sondern diese durch Tonhöhendistanzen ersetzte, da die beim Zusammenklingen erfolgende Spreizung der Intervalle berücksichtigt werden müsse.In diesem Zusammenhang stellte er folgende Thesen auf [11]:

1. *"[...], daß bei der musikalischen Wahrnehmung keineswegs die einfachen Frequenzverhältnisse als solche maßgebend sind, sondern vielmehr die Tonhöhendistanzen."*
2. *"[...], daß die Tonhöhendistanzen zwischen den ersten sechs bis acht Harmonischen durch fortgesetzte Analyse von einzelnen komplexen Tönen verschiedener Grundfrequenzen vom Gehör erlernt werden."*
3. *"[...], daß es eine feste, theoretisch ideale Intonation der Intervalle nicht gibt."*

Die Aussagen haben weitgehende Folgen, und deshalb wurde experimentell überprüft, ob sie für die Intonation von Akkorden von Bedeutung sind [12]. Der Autor benutzte sechsstimmige Akkorde, deren Struktur deutlich von dem bei den ersten sechs bis acht Harmonischen gegebenen Aufbau abweicht, die also nicht durch Lernprozesse vertraut sein können. Diese

wurden in gleichstufig temperierter und in proportionaler, den Frequenzverhältnissen entsprechender Intonation hergestellt, um die beiden Versionen miteinander zu vergleichen.

Vertrauter Akkord	c^1	g^1	c^2	e^2	g^2	c^3
Proportion	2	3	4	5	6	8
Fremdartiger Akkord	c^1	as^1	d^2	ges^2	a^2	c^3
Proportion	5	8	11	14	17	20

Alle aus benachbarten Tönen gebildeten Intervalle überschreiten die kritische Bandbreite, so daß sie also keine Grundton-Schwebungen erzeugen. Es treten aber etliche Partialtöne in eng benachbarter Lage auf, so daß der entstehende Klangeindruck sowohl in temperierter als auch in proportionaler Stimmung in hohem Grade gestört sein müßte. In temperierter Stimmung klingt der Akkord tatsächlich verschwommen, unscharf, trüb; er ist mit einer verwackelten Photographie zu vergleichen. Wird die Stimmung dagegen so justiert, daß sie den Proportionen exakt entspricht, so wirkt der Akkord klar und scharf gezeichnet; er bietet jetzt - trotz Kollision etlicher Partialtöne - den Eindruck von Geschlossenheit bei klaren Konturen. Dieser Gegensatz läßt sich in entsprechender Weise mit vielen anderen Akkordpaaren bestätigen. Bei Zusammenklängen gelten also nach unseren Untersuchungen tatsächlich die einfachen Proportionen. Die Klangqualität kann durch das Terhardtsche Modell auf der Basis der Ortstheorie nicht hinreichend erklärt werden. Es müssen daher andere Ursachen zur Erklärung der Phänomene herangezogen werden.

Die Mikrorhythmen-Theorie

Vertreter der zweiten oben erwähnten Forschungsrichtung sehen die Ursache mancher schwer erklärbaren Tatsachen auf dem Gebiet des Hörens darin, daß das Nervensystem neben dem örtlichen Code - also dem Ort der gereizten Nervenzelle - auch die zeitliche Struktur der Nervenimpulsfolgen als Informationsträger benutzt. Dieser Gedanke knüpft an eine Reihe von Hypothesen, die zur Zeit ihrer Entstehung noch nicht durch neurophysiologische Erkenntnisse gestützt werden konnten. Die Grundidee geht auf Leonhardt Euler zurück [13]. In seiner Epoche war die Wellennatur des Schalles erkannt worden, und man führte die Tonempfindung auf periodische Anstöße des Trommelfells zurück, deren Aufeinanderfolge in bestimmter Weise zeitlich geordnet ist. Dann - so folgerte Euler - müßte die mikrozeitliche Gliederung dieser Reihe von Stößen für den Organismus von Bedeutung sein.

Ein Objekt gefällt uns, heißt es im II. Kapitel des *Tentamen*, in welchem wir eine ihm innewohnende Ordnung wahrnehmen.

> *"Placent itaque ea, in quibus ordinem, qui inest, percipimus; magis autem delectabimur, si plures eiusmodi res offerantur, quarum quem continent ordinem comprehendimus; atque maximum sentiemus suavitatis gradum, si praeterea ipsarum istarum rerum ordinem, quem inter se tenent, cognoscimus."* [14]

Je leichter die Ordnung zu erfassen ist, desto vollkommener gilt es uns, und erfüllt uns deshalb mit innerer Freude.

> *"Quo facilius ordinem, qui in re proposita inest, percipimus, eo simpliciorem ac perfectionem eum existimamus idioque gaudio et laetitia quadam afficimur."* [15]

Euler stellte die Luftpulse durch Reihen von Punkten dar, deren Abstände der Periodendauer der jeweiligen Schwingung entsprechen. Bei gleichzeitig erklingenden Tönen fügen sich die Punktreihen, die die einzelnen Schallwellen repräsentieren, zu übergeordneten Gruppen zusammen, die - in Abhängigkeit vom Frequenzverhältnis - eine kleinere oder größere Zahl von Punkten umfassen und in bestimmter Weise gegliedert sind.

Ist das Frequenzverhältnis p : q, so passen in p Perioden der einen Reihe genau q Perioden der anderen. Die Ordnung in dem entstehenden Muster ist nicht von den absoluten Frequenzen

sondern allein von den Proportionszahlen abhängig. Daraus ergibt sich für Euler die Transponierbarkeit der Intervalle. Die Ordnung wird um so komplexer und damit schwerer zu erfassen, je größer die am Ausdruck des Frequenzverhältnisses beteiligten Primzahlen sind.

Oktave	2	o	o	o	o	o	o	o	o	o	o	o	o	o
	1	o		o		o		o		o		o		o
Duodezime	3	o	o	o	o	o	o	o	o	o	o	o	o	o
	1	o			o			o			o			o
Quinte	3	o	o	o	o	o	o	o	o	o	o	o	o	o
	2	o		o	o		o	o		o	o		o	o

Kurz vor der Wende zum 20. Jahrhundert nahm der Philosoph Theodor Lipps diese Gedanken wieder auf und bemühte sich - gegen den Widerstand Stumpfs - sie mit dessen Begriff "Verschmelzung" in Einklang zu bringen [16]. Lipps verlegte in seiner *Mikrorhythmen-Theorie* die Eulerschen Punktmuster in die Ebene der Nervenaktivität. Er hielt es für möglich, daß die durch Töne ausgelösten Nervenimpulsketten in den zentralen Bahnen des Hörnervensystems den Rhythmus physikalischer Schwingungen widerspiegeln.

"Nun überträgt sich freilich der Rhythmus der physikalischen Schwingungen in den physiologischen Organen in eine andere Sprache. [...] Aber die Möglichkeit besteht, [...] daß der Rhythmus dieser psychischen Vorgänge dem Rhythmus der physikalischen Schwingungen analog bleibt, so weit zum mindesten, daß das Verhältnis der psychischen Rhythmen mit dem Verhältnis der physikalischen in Vergleich gestellt werden kann." [17]

Sind die zwei Tönen zugeordneten Schwingungszahlen p und q, so werden sich auch die Perioden der durch die Schallschwingungen erzeugten nervlichen Teilvorgänge entsprechend verhalten und p Teilvorgänge des einen das gleiche Zeitintervall z ausfüllen wie q Teilvorgängen des anderen. Demnach ist die gesamte neurale Aktivität bei konsonanten Mehrklängen mikrorhythmisch gegliedert, und es wird verständlich, daß konsonante Zusammenklänge in der Wahrnehmung als geschlossene Einheiten aufgefaßt werden.

"Konsonanz besteht, sofern nach Ablauf von z immer wieder dieselben zeitlichen Zusammenordnungen eintreten. [...] Jede von mir innerlich vollzogene Weise der Zusammenordnung der Teilvorgänge erleichtert den Vollzug der folgenden gleichen Weise der Zusammenordnung." [18]

Kritiker der Mikrorhythmen-Theorie argumentierten mit dem Einwand, daß wir gerade bei den tiefsten Tönen, wo wir den Schwingungsrhythmus noch wahrnehmen können, die konsonanten Intervalle keineswegs angenehmer finden als die dissonanten. Lipps' Antwort lautete: Die psychische Zusammenordnung ist um so vollkommener, je rascher sie erfolgt, d.h. je kürzer z ist. Bei den tiefsten Tönen aber ist z um ein Vielfaches länger als in den höheren Lagen, und mit der Länge des Zeitintervalls z nimmt die bindende Kraft ab.

Auch im Falle leicht verstimmter Konsonanzen, wie sie bei der temperierten Stimmung vorliegen, kann dieses Bild gelten und das sogenannte Zurechthören erklären. Sind z.B. zwei Frequenzen p = 100 Hz und q = 201 Hz, die Periodendauern also T_p = 10 ms und T_q = 4,9751 ms, so ist zwar das Zeitintervall, das bis zum exakten Zusammentreffen der beiden Wellenzüge vergeht, sehr lang. Zugleich aber hat sich der Unterschied zwischen den einzelnen Teilperioden vermindert; eine Periode von 10 ms ist annähernd gleichlang wie 2 mal 4,9751 ms. Auch dann, wenn die Periodendauern zweier Impulssequenzen nur annähernd übereinstimmen, könnte dies innerhalb gewisser Grenzen wie völlige Übereinstimmung aufgefaßt werden, so wie im optischen etwa ein annähernd kreisförmiges oder quadratisches Gebilde wie die regelmäßige Form wirken kann, sofern die Annäherung genügend groß ist.

Die Mikrorhythmen-Theorie geht davon aus, daß die Periodizität des Schalles in der Nervenaktion erhalten bleibt. Dies wurde zunächst grundsätzlich bestritten. Zwar gelangte man hinsichtlich der in den Nerven stattfindenden "Teilvorgänge", von denen Lipps gesprochen hatte, um die Jahrhundertwende zu neuen Erkenntnissen. Es wurde ein wesentliches Prinzip der Nervenfunktion entdeckt, nämlich ihre diskontinuierliche Aktion, das sogenannte Alles-oder-Nichts-Prinzip. Die Nervenzellen arbeiten relaisartig, jeweils bei Überschreitung einer Reizschwelle wird

ein Nervenimpuls ausgelöst. Anschließend benötigt die Zelle ein kurzes Zeitintervall (Refraktärperiode), um in den reaktionsbereiten Zustand zurückzukehren. Die Impulse in den Fasern des von der Cochlea zum Gehirn führenden Hörnerven könnten also durch periodische Schallvorgänge synchronisiert werden, so daß die Abstände zwischen den einzelnen Nervenimpulsen die Periodendauer der Schallwelle abbilden. Auf dieser neurophysiologischen Basis könnten die von Lipps angenommenen Prozesse ablaufen. Die Refraktärperiode setzt der Synchronisierung der Nervenimpulse jedoch eine Grenze. Galambos und Davis fanden am Hörnerven der Katze, daß eine einzelne Faser kurzzeitig bis zu 400, nach Adaptation nur noch bis zu 200 Impulse pro Sekunde gibt [19].

Im Gegensatz dazu hatten Wever und Bray an Faserbündeln gefunden, daß die Grenze für die Synchronisierung dieser summierten Potentiale zwischen 3000 und 4000 Hz liegt [20]. Spätere Untersuchungen von Davis bestätigten das. Da einzelne Fasern die hohen beobachteten Frequenzen nicht übertragen können, deuteten Wever und Bray die Beobachtung dahin, daß sie das Resultat der Kombination vieler Fasern zu gemeinsamer Aktion ist. Wever ging davon aus, daß auch bei dem einfachsten akustischen Reiz jeweils eine große Anzahl von Nervenzellen der Erregung ausgesetzt ist, aber dennoch nicht alle Zellen gleichzeitig reagieren. Die einzelnen Zelle unterliegen einer stoffwechselbedingten zyklischen Variation ihrer Reaktionsbereitschaft. Durch diese Unterschiede wird die Gesamtgruppe in verschiedene Teilgruppen aufgeteilt, die zu verschiedenen Zeitpunkten auf den Schallreiz antworten. Der daraus resultierende Vorgang der Nervenaktion kann verglichen werden mit dem abwechselnden Schlagen einer Trommel mit zwei Stöcken [21]. Diese alternierende Aktion innerhalb einer Gruppe von Nervenzellen nannte Wever *volley-principle* [22].

Erregung durch den Schallvorgang:	o	o	o	o	o	o	o	o
Reaktion von Zellgruppe 1:	*		*		*		*	
Reaktion von Zellgruppe 2:		*		*		*		*
Impulsfolge im Faserbündel:	*	*	*	*	*	*	*	*

Die Tonhöhenempfindung hat damit zwei physiologische Korrelate. Der Schallreiz wird in der Cochlea einerseits in Komponenten aufgeteilt, die an verschiedenen Stellen des Cortischen Organs zur Wirkung kommen. Außerdem wird mit Hilfe des Volley-Prinzips die Zeitstruktur des Reizes im Muster der Nervenimpulse repräsentiert. Lothar Cremer hatte darauf aufmerksam gemacht, daß dies schon aus physikalischen Gründen erforderlich sei [23]:

"Die klassische Theorie der Tonempfindungen beschäftigte sich zunächst nur mit den stationären Klängen. Sobald wir aber die ständig wechselnden spektralen Verteilungen in Sprache und Musik erfassen wollen, entsteht die Frage der Einschwingzeit des Analysators, mit dieser aber unumgänglich verbunden ist die seiner Analysierschärfe." [24]

Obgleich noch viele Jahrzehnte vergingen, bis alle neurologischen Bedingungen dafür, daß das Lipps'sche Modell der Wirklichkeit entsprechen könnte, erforscht waren, war Erich Moritz von Hornbostel, eine der überragenden Persönlichkeiten in der Musikwissenschaft unseres Jahrhunderts, bereits in den zwanziger Jahren von Lipps' Idee überzeugt und entwickelte dessen Gedankengänge weiter [25]. Er stellte die Hypothese auf, daß in höheren Zentren des Gehirns auch Nervenimpulse, die in gleichen Abständen aufeinanderfolgen, zu übergeordneten Perioden zusammengefaßt werden.

"Man kann eine gleichmäßige Folge von Lichtblitzen nicht sehen, eine gleichmäßige Klopfreihe nicht hören oder ausführen, ohne sie unwillkürlich und oft unwissentlich zu gliedern. Wodurch wir gliedern, [...] ist gleichgültig. Wesentlich ist nur, daß die vielen Einzelnen sich zu Grüppchen, diese zu Gruppen zusammenschließen, das Ganze sich ordnet und gestaltet und so überschaubarer, faßbarer wird. [...] Wir wollen nun annehmen, das Phänomen der Tonigkeit beruhe auf einer solchen gliedernden Gestaltung der an sich gleichmäßig periodischen zentral-physiologischen Vorgänge." [26]

Der Tonigkeit, der ja ein unmittelbares physikalisches Korrelat fehlt, wird in der Zeitstruktur der Nervenimpulsfolge ein sekundäres physiologisches zugewiesen. Sobald die Schwingungsform der Luft ungleichmäßig wird, zerfließt die rhythmische Ordnung, die Tonigkeit verschwindet, und der Klang wird geräuschhaft. Ähnliche Verhältnisse scheinen bei Amusischen zu bestehen: bei ihnen dürfte nach Hornbostel die Fähigkeit zur Gliederung der Impulsketten mangelhaft sein oder

ganz fehlen. Auch für Musikalische tritt die Tonigkeit bei sehr hohen, sehr tiefen und sehr kurzen Tönen stark zurück. Sie sei, schreibt Hornbostel, als das labilste, entwicklungsgeschichtlich jüngste Moment an den akustischen Erscheinungen anzusehen.

Das Modell kann auch auf Tonfolgen und Zusammenklänge angewandt werden. Wenn die einzelnen Nervenfaserbündel Impulsfolgen zu bestimmten Integrationszentren leiten, wo das Zusammentreffen der Impulse aus verschiedenen Faserbündeln registriert wird, so läßt sich auf dieser Basis auch das Phänomen der Tonverwandtschaft verstehen.

"Dem phänomenalen Zusammenpassen wird ein Zusammenpassen der Strukturen der zentral-physiologischen Prozesse, der Einheitlichkeit der Gesamterscheinung die Einheitlichkeit der physiologischen Gesamtstruktur entsprechen." [27]

Tonverwandtschaft besteht nach Hornbostels Vorstellung in dem Maße, wie sich die Nervenimpulsfolgen nach einem einheitlichen Gliederungsprinzip ordnen lassen. Ein erkling-ender Ton erzeugt eine Nervenimpulsfolge, und setzt durch deren zeitliche Gliederung eine Art Taktschema, auf das die durch den folgenden Ton ausgelöste Nervenimpulsfolge bezogen wird. Erklingt nach einem Ton dessen Oktave, so bleibt die Gliederung dieselbe, es wechselt nur die Anzahl der zu einem Glied zusammengeschlossenen Unterglieder.

```
o       o       o       o       o
o   o   o   o   o   o   o   o   o
```

Passen die Gliederungen dagegen nicht zusammen, so schlägt das Gliederungsschema um.

```
o     o     o     o     o
o     o     o     o     o     o
```

Das Bild von der Struktur der zentral-physiologischen Prozesse wird weiterhin in Feldstruktur und Teilstrukturen differenzierte. Alle Strukturen gelten - wie bei Euler - als um so näher verwandt, je niedriger die Primzahlen sind, auf die sie sich gründen. Hornbostel stellte folgende Regeln zur Bestimmung des Verwandtschaftsgrades von Tönen auf:

1. Als nächstverwandt müssen Strukturen mit gleichem und bezüglich Hauptgliederung (Feldstruktur) und Untergliederung (Teilstruktur) einheitlichem Gliederungsprinzip gelten.
2. Strukturen sind um so näher verwandt, je niedriger die Primzahlen sind, nach denen die Gliederung aufgebaut ist, also nach 2^n, 3^n usw.
3. Die Feldstruktur hat für die Verwandtschaft größeres Gewicht als jedeTeilstruktur.

```
Feldstruktur    o                   o                   o                   o                   o
Teilstruktur 2  o         o         o         o         o         o         o         o         o
Teilstruktur 3  o     o     o     o     o     o     o     o     o     o     o     o     o
```

Der Ton, der die Primzahl 2 in die Feldstruktur einbringt, hat das größte Gewicht für die Gesamtstruktur. Der dadurch entstehende Zusammenhang, so postuliert die Hornbostelsche Strukturtheorie, wird als Tonalität bewußt. Diese Auffassung deckt sich mit Max F. Meyers Theorie der Melodik [28]. Die Intervalle werden hier durch Frequenzverhältnisse repräsentiert, und das von Meyer sogenannte *Lipps-Meyer-Gesetz* besagt, daß der Entspannung ausdrückende Grundton einer Melodie in dieser Darstellung stets eine Potenz von 2 ist. Dies sei am Beispiel einiger Tonschritte gezeigt:

d^1 - c^1	e^1 - c^1	g - c^1	g - c	h - c^1
9 : 8	5 : 4	3 : 4	3 : 2	15 : 16

Ergebnisse der neueren Hörforschung

Seit Mitte unseres Jahrhunderts hat die neurophysiologische Forschung nach und nach immer mehr Beweise dafür gefunden, daß die Zeitstruktur des Schalles in Gestalt reizsynchroner Nervenimpulse im unteren Teil des Hörnervensystems tatsächlich erhalten bleibt [29].

"Die als Folge des Schallsignals sich hin und her bewegenden Stereozilien der Sinneszellen führen nur bei Displazierung in einer Richtung zu Depolarisation und Freisetzung von Transmittern. […] Die Häufigkeit der Phasen entspricht naturgemäß der Frequenz des Schallsignals. […] Die Entladungsraten spiegeln daher das Zeitmuster des Schallsignals wider. Dieser Mechanismus funktioniert bis etwa 5 kHz. Die Periodizitätsanalyse wird von kochleären Implantaten genutzt, welche […] in der Lage sind, trotz Reizung an nur einem Ort einen gewissen Frequenzumfang zu übertragen." [30]

Damit sind die lange geleugneten Voraussetzungen für die Lipps'sche Mikrorhythmen-Theorie gegeben, und die gehörpsychologischen Befunde sprechen dafür, daß die in diesem zeitlichen Muster verschlüsselte Information in derjenigen Weise oder ähnlich ausgewertet wird, wie Lipps es postuliert hatte.

Der Verfasser hatte im Rahmen einer kritischen Analyse der klassischen Hörtheorien eine Reihe von Tonhöhenwahrnehmungen zusammengestellt, die allein auf der Auswertung von periodischen Gliederungen der Schallsignale, nicht aber durch die *Ortstheorie* erklärt werden können und dafür sprechen, daß Zeitintervalle im Nervenimpulsmuster bei der Tonhöhen-erkennung eine wichtige Rolle spielen müssen [31].

Offensichtlich werden Zeitintervalle im Nervensystem mit großer Präzision übertragen. Der Richtungseindruck beim Hören entsteht - neben der Auswertung von Intensitätsunterschieden, die zwischen beiden Ohren entstehen - auf Grund der äußerst geringen Zeitdifferenz, mit der die Schallwellen die beiden Ohren erreichen. Die Laufzeitdifferenzen betragen nur Bruchteile von Millisekunden; bei der eben erkennbaren Abweichung von der Medianebene beträgt die Zeitdifferenz um 30 Mikrosekunden [32]. Diese kann nur erfaßt und ausgewertet werden, wenn die zeitliche Struktur des Schalles in den von beiden Ohren kommenden Nervenbahnen zumindest bis zu den Bahnkreuzungen im Hirnstamm sehr genau repräsentiert ist.

Einen weiteren Beweis dafür, daß die Periodizität des Schalles tatsächlich in der Nervenaktion erhalten bleibt, liefert das Phänomen der binauralen Schwebungen. Werden beide Ohren getrennt mit zwei Schallwellen gereizt, die bei monauraler Darbietung sehr langsame Schwebungen ergeben würden, so scheint der Ton zwischen beiden Seiten des Kopfes hin- und herzuschwingen oder um den Kopf herumzuwandern. In der deutschsprachigen Literatur wurde daher von Drehtönen gesprochen. Dies ist eine dynamische Form der gewöhnlichen, statischen Lokalisation nach dem Phasenverhältnis. Die Erscheinung wurde bereits in der zweiten Hälfte des vorigen Jahrhunderts beschrieben. Die meisten Wissenschaftler glaubten jedoch, daß es sich um monaurale, durch Knochenleitung entstandene Schwebungen handele. Durch moderne Versuche unter einwandfreien technischen Bedingungen konnten sie jedoch zweifelsfrei nachgewiesen werden. Joseph Matzker hat sie eingehend untersucht und auf dieser Basis einen Test zum Nachweis zerebraler Hörstörungen entwickelt [33]. Es steht fest, daß es sich um Wechselwirkungen innerhalb des Nervensystems handelt, die durch die kontralaterale Verschaltung der Nervenbahnen ermöglicht werden.

Wenn aber kontralaterale Effekte bewiesen sind, so können auch unilaterale Wechselwirkungen zwischen den an verschiedenen Orten der Basilarmembran ausgelösten Nervenimpulsen angenommen werden. Es gibt eine Anzahl von Beobachtungen, die auf Grund eines derartigen wechselseitigen Zusammenwirkens zu verstehen wären. Erklingen zwei Vokale gleichzeitig, so überlagern sich deren Formantbereiche und bilden ein sehr komplexes Spektrum; dennoch kann man die beiden Vokale identifizieren. Broadbent und Ladefoged fanden, daß Vokale auch dann identifiziert werden, wenn man das gesamte Spektrum in der Weise aufteilt, daß der Unterformant dem einen Ohr und der Oberformant dem anderen Ohr zugeleitet wird. Eine Verschmelzung findet jedoch nicht statt, wenn die beiden Formantbänder auf verschiedene Grundfrequenzen bezogen sind, also unterschiedliche Hüllkurvengliederungen aufweisen [34].

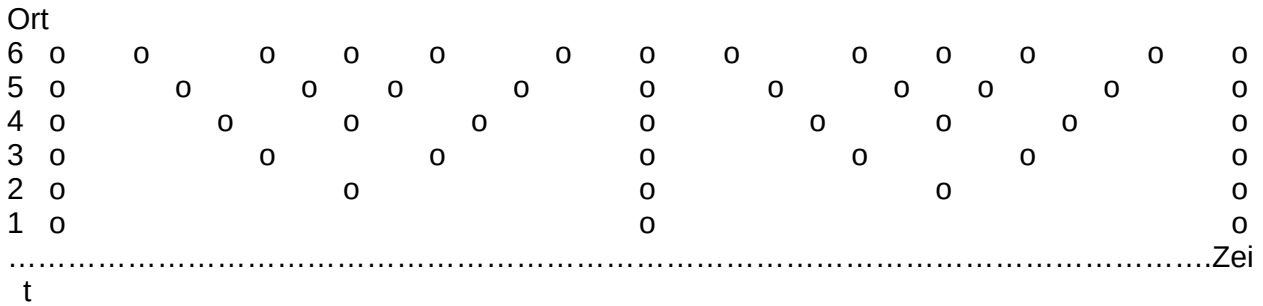
Bei einem Vokal sind alle Komponenten des Spektrums Vielfache der gleichen Grund-frequenz, da sie durch die Stimmlippenschwingung synchronisiert werden. Diese Bedingung wird in entsprechender Weise von allen periodischen Schallschwingungen erfüllt. Ein komplex-periodischer Schall erzeugt im Innenohr eine örtlich aufgefächerte Erregungsverteilung. Der an den verschiedenen Orten des Cortischen Organs erzeugte Erregungsverlauf folgt den Druckphasen der jeweils zugeordneten Komponente der Schallwelle. In einer festen Phase des Erregungszyklus werden Nervenimpulse ausgelöst, deren zeitlicher Abstand die Periodendauer der Partialschwingung repräsentiert. Die Frequenzen der in den verschiedenen Regionen

ausgelösten Nervenimpulssalven stehen also in einfachen Verhältnissen zueinander. Mit Hilfe des Salvenprinzips können Gruppen von Nervenzellen diese Frequenzen bis etwa 4 - 5 kHz übertragen.

Das Mikrorhythmen-Modell

Rufen wir nun das hypothetische Modell in Erinnerung, das Lipps aufgestellt hatte, und deuten es auf Grund der geschilderten Erkenntnisse. Die Senkrechte repräsentiert in dem Modell die Orte des Cortischen Organs, an denen durch die Partialtöne Impulse ausgelöst werden, die Horizontale bildet die zeitliche Folge der an den einzelnen Orten ausgelösten Nervenimpulse ab. Natürlich ist das Schema nicht als eine topologisch getreue Abbildung zu verstehen, aber es symbolisiert die bereits durch einen Einzelton ausgelöste Erregungsverteilung. Insofern gehen wir über Lipps hinaus und stellen die These auf:

Die Tonhöhenwahrnehmung beruht auf einem Mustererkennungsprozeß, der die örtliche und zeitliche Gliederung des Nervenimpulsmusters auswertet.



Örtlich und zeitlich gegliedertes Nervenimpulsmuster

Man sieht, daß die an den einzelnen Orten der Basilarmembran ausgelösten Impulse in regelmäßigen Zeitabständen zusammentreffen. Dieses Zeitintervall ist die Periodendauer der Gesamtschwingung. Die Periodizität ist im Nervenimpulsmuster um so deutlicher ausgeprägt, je mehr Komponenten zusammenwirken. Die von Hornbostel vermutete "gliedernde Gestaltung der zentral-physiologischen Vorgänge" ist also vorhanden, ohne daß dafür eine psychische Tätigkeit notwendig wäre. Die Gliederung könnte mit Hilfe von Koinzidenz-Detektoren in frühen Zentren der Hörbahn - möglicherweise in den Cochleariskernen im Hirnstamm - registriert werden. Da die qualitative Komponente der Tonhöhe - das Chroma [35] - nur dann wahrgenommen wird, wenn die Schallschwingung periodisch ist, sehen wir die zeitliche Gliederung des Erregungsmusters als physiologische Basis des Chromas an, die örtliche Verteilung der Erregung gilt uns als Korrelat der Tonhelligkeit. Oberhalb von 4000 Hz verblaßt die Chromaqualität und verschwindet schließlich völlig, da die Schallfrequenzen in diesem Bereich auch mittels des Salvenprinzips nicht mehr frequenzsynchron übertragen werden können.

Die Qualität Chroma basiert in diesem Erklärungsmodell weder auf der Eigenschaft eines Elementes (der Grundfrequenz) noch auf der Summe von Elementareigenschaften (Teiltöne bzw. Tonigkeiten), sondern auf der Gesamtheit der Relationen, in denen die Elemente des Musters, nämlich die Nervenimpulse, untereinander stehen. Auch das Chroma ist demnach eine Gestaltqualität. Die Partialschwingungen verstärken die Ausprägung der zeitlichen Gliederung des Musters, unabhängig von ihrer individuellen Periodendauer. Jeder Musiker weiß, daß die Präzision des Tonhöhenurteils bei hellen, obertonreichen Tönen größer ist als bei obertonarmen. Unter der gleichen Bedingung steigt die Urteilssicherheit bei Absoluthörern. Dieses bestätigt die Chroma-Hypothese.

Die periodische Gliederung des Nervenimpulsmusters - die Voraussetzung für die Chroma-Empfindung - bleibt mit gleicher Periodendauer auch dann erhalten und ohne weiteres erkennbar, wenn die Grundfrequenz fehlt und die Tonhöhe als Residuum hörbar bleibt. Zur Erklärung dieses Phänomens werden also keine Hilfshypothesen benötigt. Außerdem wird klar - und darüber ist sehr kontrovers diskutiert worden -, warum und wie die Einzelelemente des Musters das Chroma des Grundtones verstärken können, ohne dabei ihre Individualität zu verlieren wie etwa zusammenfließende Farben. Sie bleiben mit ihrer individuellen Periodendauer innerhalb des Musters präsent. Wird die Aufmerksamkeit auf eine einzelne Impulskette gelenkt, so kann diese Teilfrequenz identifiziert, also ein Oberton herausgehört werden. Das Erkennen des Gesamtmusters ist allerdings die einfachere Leistung. Die Erfahrung bestätigt dies: Jeder

Gesunde hört die Grundtonhöhe ohne Mühe, während das Heraushören von Obertönen besondere Konzentration und einige Übung erfordert.

Das Modell der neuralen Muster darf nur dann Glaubwürdigkeit für sich in Anspruch nehmen, wenn es in biologischem Sinne als zweckmäßig gelten kann. Hören ist eine Tätigkeit, die - biologisch gesehen - auf das Erkennen und Bewerten flüchtiger Laute gerichtet ist, die in Form komplexer Schallwellen auf unser Ohr treffen. Bedeutungsvolle Schallwellen aber gelangen selten isoliert zum Hörorgan, sondern sind eingebettet in das Gemisch des übrigen von der jeweiligen Umwelt erzeugten Schalles. Man stelle sich vor, das Gehör würde stimmhafte Sprachlaute dadurch identifizieren, daß es eine große Anzahl von Orten auf der Basilarmembran, die Spektrallinien entsprechen, hinsichtlich ihrer Abstände auswertet. Versetzt man sich nun in eine lebhafte Gesellschaft, wo viele Worte durcheinander erklingen, so wäre das Ohr in einer hoffnungslosen Situation.

Die zusammengehörenden, ein Signal bildenden Schwingungskomponenten von unterschiedlicher Frequenz haben aber ein Merkmal gemeinsam, nämlich, daß sie durch den mechanischen Prozeß ihrer Erzeugung - durch den Rhythmus der Stimmlippenschwingung - streng phasengekoppelt sind. Die Frequenz jeder Partialschwingung bildet daher ein ganzzahliges Vielfaches der Stimmlippenfrequenz. Im Innenohr wird das Frequenzgemisch aufgefächert und es werden in einem ausgedehnten Abschnitt des Cortischen Organs Nervenimpulse ausgelöst. Wenn die Impulsketten aller erregten Nervenzellen in der nächsten Station der Hörbahn aufeinander bezogen werden, so heben sich diejenigen Nervenfasern aus der Gesamtheit heraus, bei denen die Impulse in regelmäßigen Zeitintervallen gleichzeitig eintreffen.

Die Gleichzeitigkeit führt zum Zusammenschluß der Aktivität dieser Nervenfasern zu einem Muster, das sich von der übrigen, nicht synchronisierten Nervenaktivität abhebt. Bei diesem Übertragungsverfahren wird die Sicherheit des Signalempfangs dadurch erhöht, daß die Information von einer Gruppe von Nervenfasern, die gemeinsam das Muster übertragen, zu den zentralen Verarbeitungsstellen im Gehirn geleitet wird. Die Auswertung eines derartigen, hier vom Prinzip her angedeuteten Musters wird als Zeitreihen-Korrelationsanalyse bezeichnet. Sie ermöglicht die biologisch sinnvolle Fähigkeit, zusammengehörige Information zu verbinden, die dann als Einheit wahrgenommen werden kann. Licklider erkannte als erster, daß eine Zeitreihen-Korrelationsanalyse Hörphänomene wie Perioden- und Pulstöne erklären würde, und daß die dafür notwendigen anatomisch-physiologischen Voraussetzungen durch Neuronennetze erfüllt werden können [36].

Nach dem gleichen Prinzip läßt sich auch das Verschmelzen konsonanter Zusammenklänge erklären. Auf Grund zahlreicher Beobachtungen hat der Verfasser dieses Phänomen in der genannten Schrift auf die zeitliche *Kohärenz* des Nervenimpulsmusters, d.h. auf den taktmäßigen Zusammenhang der ausgelösten Nervenimpulse, zurückgeführt [37]. Dieser zeitliche "Gleichschritt" hätte zur Folge, daß die einzelnen Komponenten eines konsonanten Zusammenklanges miteinander verschmelzen, was in der Wahrnehmung als Klangverwandtschaft in Erscheinung tritt. Die Klangverwandtschaft ist also auf Grund der Informationsverarbeitung im Nervensystem naturgegeben und nicht aufgrund von Lernprozessen entstanden. Auch die allgemein vertraute Erfahrung, daß ein Wechsel der Klangfarbe nur relativ geringen Einfluß auf die Intervallqualität hat, findet ihre Erklärung; denn auch bei Wegfall oder Hinzufügung einer Anzahl von Partialtönen bleibt die Feldstruktur des Impulsmusters erhalten, so daß die Kohärenz der Impulsmuster sich nicht grundsätzlich verändert.

Daß die Sonanz von Intervallen außerdem auch durch Obertonkonstellationen beeinflußt werden kann, wird dadurch nicht ignoriert. Die Hörerfahrung, daß Klänge sich auf Grund ihres Spektrums unterschiedlich gut zu Mehrklängen zusammenfügen lassen, brachte Arnold Schering in dem Begriffspaar *Spaltklang -- Verschmelzungsklang* zum Ausdruck [38]. Wenn starke Anteile der beteiligten Schallwellen in kritischen Bändern zusammentreffenden, so hat dies eine Diskontinuität der Impulssalven und damit das Entstehen von Rauhigkeit zur Folge. Auf Grund dieses Kriteriums lassen sich Spaltklänge, die sich durch ihre Rauhigkeit der Verschmelzung widersetzen, von Verschmelzungsklängen unterscheiden, die durch die zugrundeliegende Schwingungsstruktur wenig Anlaß für "Störungen des Zusammenklanges" (Helmholtz) geben.

Abschließend sei ein weiteres Phänomen geschildert, das sich ebenfalls in das Gesamtbild fügt. Seit langem ist bekannt, daß bei Abweichung von einfachen Frequenzverhältnissen auch dann Schwebungen zu beobachten sind, wenn man durch Sinuswellen erzeugte sogenannte

Sinustöne verwendet, wenn also keine Obertöne für die Erklärung der Schwebungen vorhanden sind. Man hat die Beobachtung früher so interpretiert, daß man annahm, die Schallwellen würden bei der Übertragung ins Innenohr nichtlinear verzerrt. Eine verzerrte Schwingung ist - wie jede komplex-periodische Funktion - durch eine Fourieranalyse in eine Reihe von Sinusschwingungen aufzulösen. Die während des Übertragungsvorganges entstehenden "subjektiven" Obertöne würden nun an die Stelle der normalerweise vorhandenen objektiven Partialtöne treten und u.a. auch Schwebungen erzeugen. Zwar treten bei hohen Intensitäten tatsächlich Effekte nichtlinearer Übertragung auf, aber die genannten Schwebungen sind auch bei niedriger Intensität zu beobachten, bei der die Schallwelle erwiesenermaßen nicht verzerrt wird [39]. Dennoch hat sich dieser Erklärungsversuch hartnäckig gehalten.

Die Vertreter der Nichtlinearitäts-Hypothese führen die Schwebungen bei Sinustönen entweder auf im Ohr entstandene Obertöne oder Differenztöne zurück. Diese Hypothesen aber wurden durch die grundlegenden Arbeiten Plomps widerlegt [40]. Plomp bildete verstimmte Intervalle aus Sinustönen und verdeckte die Tonhöhenbereiche, in denen die hypothetischen Ober- bzw. Differenztöne liegen müßten, durch Rauschbänder. Es zeigte sich jedoch, daß die Schwebungen dadurch nicht verschwinden, daß das Konzept der kritischen Bandbreite in diesem Falle also nicht herangezogen werden kann. Roederer bezeichnet diese Schwebungen als *Schwebungen zweiter Ordnung* [41]. Diese Schwebungen unterscheiden sich von denjenigen der ersten Ordnung dadurch, daß man bei langsamen Schwebungen, die man im einzelnen verfolgen kann, nicht periodische Lautstärkeschwankungen sondern eine Fluktuation des Klangcharakters empfindet. Vergrößert man den Grad der Verstimmung, so verschmelzen die Schwebungen wie diejenigen erster Ordnung zu einem Gesamteindruck: es entsteht ein verschwommener, getrübter Klang.

Genau die gleiche Beobachtung machten wir in dem geschilderten Experiment mit den vielstimmigen Akkorden, wenn diese in temperierter Stimmung erklangen. Wir vermuten daher, daß die Schwebungen zweiter Ordnung generell eine weit größere Bedeutung für den Klangeindruck haben, als bisher angenommen wurde. Beziehen wir die Beobachtung auf die Hypothese des neuralen Mustererkennungsprozesses: Wenn die Frequenzen zweier oder mehrerer Schallwellen um einen geringen Betrag vom Verhältnis kleiner ganzer Zahlen abweichen, so verschieben sich die Schwingungsphasen der Wellen allmählich gegeneinander. Wird der Schall in der Cochlea in Komponenten analysiert, die an verschiedenen Orten des Cortischen Organs jeweils in einer bestimmten Schwingungsphase Nervenimpulse auslösen, so haben die Phasenverschiebungen zur Folge, daß die Impulsfolgen nicht fest synchronisiert sind sondern sich periodisch gegeneinander verschieben. Das Orts-Zeit-Muster der Nervenimpulse fluktuiert, der Klangeindruck - so beschrieben wir es oben - verschwimmt wie die Konturen einer verwackelten Photographie.

Die verschiedenen Beobachtungen bestätigen die hinsichtlich der Auswertung von Nervenimpulsmustern aufgestellten Hypothesen und erhöhen dadurch die Glaubwürdigkeit einer Vorstellung vom Prozeß des Hörens, in der die Ideen von Leonhard Euler, Theodor Lipps und Erich Moritz von Hornbostel lebendig sind.

Anmerkungen

[1] Chr. Freiherr von Ehrenfels: *Über "Gestaltqualitäten"*, in: Vierteljahresschrift für wissenschaftliche Philosophie 14 (1890) S. 249-292.
[2] E. Mach: *Beiträge zur Analyse der Empfindungen*, Jena 1886.
[3] C. Stumpf: *Tonpsychologie*, Bd. 2, Leipzig 1890.
[4] ebd. S. 213.
[5] H. v. Helmholtz: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Braunschweig 1863, ^{6}1913.
[6] G. v. Békésy: *Experiments in hearing*, New York 1960.
[7] J. F. Schouten: *The residue, a new component in subjective sound analysis*, in: Proc. Nederlandsche Akademie van Wetenschappen 45 (1940) S. 256-365.
[8] E. Terhardt: *Zur Tonhöhenwahrnehmung von Klängen*, in: Acustica 26 (1972), S. 173-186.

[9]E. Terhardt: *Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz*, in: Acustica 36 (1976/77), S. 121-137.
[10] H. v. Helmholtz, [6]1913, S. 320.
[11] E. Terhardt, 1976/77, S. 129, 132, 135.
[12] H.-P. Hesse: *Zum Konsonanzempfinden bei vielstimmigen Akkorden*, in: Internationaler Musikwissenschaftlicher Kongreß zum Mozartjahr 1991, Bericht (hg. v. I. Fuchs), Tutzing 1993, S. 669-675.
[13] L. Euler: *Tentamen novae theoriae musicae*, St. Petersburg 1739.
[14] ebd. II S. 12.
[15] ebd. II S. 13
[16] Th. Lipps: *Tonverwandtschaft und Tonverschmelzung*, in: Zeitschrift für Psychologie und Physiologie der Sinnesorgane 19 (1899), S. 1-40.
[17] ebd. S. 29.
[18] ebd. S. 32.
[19] R. Galambos & H. Davis: *The response of single auditory-nerve fibers to acoustic stimulation*, in: Journ. Neurophysiol. 7 (1943), S. 39-57.
[20] E. G. Wever & C. W. Bray: *Action currents in the auditory nerve in response to acoustical stimulation*, in: Proc. Nat. Acad. Sci. 16 (1930), S. 344-350.
[21] Natürlich ist dies Bild schematisch vereinfacht. Das Abwechseln innerhalb der Gesamtgruppe folgt in Wirklichkeit einer stochastischen Verteilung.
[22] E. G. Wever: *Theory of hearing*, New York [2]1957.
[23] L. Cremer: *Über die ungelösten Probleme in der Theorie der Tonempfindungen*, in: Acustica 1 (1951) S. 83-96.
[24] ebd. S. 87. Es besteht zwischen beiden Unschärfen die Relation t x f = 1.
[25] E. M. v. Hornbostel: *Psychologie der Gehörserscheinungen*, in: Handbuch der normalen u. pathologischen Physiologie 11 (1926) S. 701-730.
[26] ebd. S. 713.
Hornbostel bezeichnete den linear steigenden Tonhöhenaspekt als *Helligkeit* und gab der zyklischen, im Oktavabstand wiederkehrenden Toneigenschaft den Namen *Tonigkeit.*
[27] ebd. S. 718.
[28] M. F. Meyer: *Contributions to a psychological theory of music*, Columbia, Mo. 1901.
[29] I. Tasaki: *Nerve impulses in individual auditory nerve fibers of Guinea pig*, in: Journ. Neurophysiol. 17 (1954), S. 97-122.
E. F. Evans: *Central mechanisms relevant to the neural analysis of simple and complex sounds*, in: O.-J. Grüsser & R. Klinke (Hg.): Zeichenerkennung durch biologische und technische Systeme, Berlin u.a. 1971, S. 328-343.
E. F. Evans: *Pitch and cochlear nerve fibre temporal discharge patterns*, in: R. Klinke & R. Hartmann (Hg.): Hearing - Physiological bases and psychophysics, Berlin u.a. 1983, S. 140-146.
G. Langner: *Evidence for neuronal periodicity detection in the auditory system of theGuinea fowl: Implications for pitch analysis in the time domain,* in: Experimental Brain Research 52 (1983) S. 333-355.
[30] H.-P. Zenner: *Hören - Physiologie, Biochemie, Zell- und Neurobiologie*, Stuttgart; New York 1994, S.27 f.
[31] H.-P. Hesse: *Die Wahrnehmung von Tonhöhe und Klangfarbe als Problem der Hörtheorie*, Köln 1972.
[32] J. Blauert: *Räumliches Hören*, Stuttgart 1974.
[33] J. Matzker: *Ein binauraler Hörsynthese-Test zum Nachweis zerebraler Hörstörungen*, Stuttgart 1958.
[34] D. E. Broadbent & P. Ladefoged: *On the fusion of sounds reaching different sense organs*, in: JASA 29 (1957) S. 708-710.
[35] A. Bachem hat anstelle des von Hornbostel benutzten Begriffs *Tonigkeit* den Ausdruck *Chroma* eingeführt; beide Worte sind synonym.
[36] J. C. R. Licklider: *A duplex theory of pitch perception*, in: Experientia 7 (1951)
S. 128-134.
[37] H.-P. Hesse, 1972, S. 154.
[38] A Schering: *Historische und nationale Klangstile*, in: Jahrbuch Peters 34 (1927)

S. 31-44.
[39] J. G. Roederer: *Physikalische und psychoakustische Grundlagen der Musik*, Berlin; Heidelberg; New York 1977, S. 37 f.
[40] R. Plomp: *Experiments on tone perception*, Soesterberg 1966.
R. Plomp: *Beats of mistuned consonances*, in: JASA 42 (1967) S. 462-474.
[41] J. G. Roederer, 1977, S. 42 f.