



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

**Residualton & Neuronale Autokorrelation
in Modellen
der Tonhöhenwahrnehmung**

Verfasserin

Bettina Wimmer

angestrebter akademischer Grad

Magistra der Philosophie (Mag.phil.)

Wien, 2012

Studienkennzahl lt. Studienblatt:
Studienrichtung lt. Studienblatt:
Betreuer:

A 316
Musikwissenschaft
Univ.-Prof. Dr. Christoph Reuter

„I must admit that I have never liked to work hard [...].

But there is one thing I am always ready to do,
and that is to look at beautiful things.“

Georg von Békésy in seiner Nobel Lecture, 1961

Vorwort

Während meines Studiums war ich lange Zeit davon überzeugt, dass ich meine Diplomarbeit über zeitgenössische Musik schreiben möchte. Nach und nach gewann jedoch mein Interesse an der Systematischen Musikwissenschaft die Oberhand. Daher trat ich 2009 an Univ.-Prof. Dr. Christoph Reuter heran, mit dem Wunsch nach einem vielseitigen Systematik-Diplomarbeitsthema, das es mir ermöglicht meine im Laufe des Studiums erworbenen Musikwissenschafts-Kenntnisse mit meinem Interesse an Mathematik und Naturwissenschaften zu verbinden. Daraufhin schlug Prof. Reuter mir vor, mich mit dem Thema Residualton auseinanderzusetzen. Ein Phänomen, das zunächst in der Musik bekannt wurde und inzwischen auf biologischer und neurologischer Ebene erforscht wird.

Im Laufe der vergangenen drei Jahre habe ich mich nun, neben meinem Mathematik-Studium, intensiv mit dem Thema beschäftigt und bald erkannt, dass man als Musikwissenschaftler erst nach und nach mit der naturwissenschaftlichen Terminologie vertraut wird. Auf Grund der boomenden Forschung im Bereich des menschlichen Gehirns sind die Anzahl der Artikel und Publikationen zu diesem Fachgebiet in den letzten Jahren geradezu explodiert. Insofern ist es mir nicht leicht gefallen, eine Auswahl zu treffen, um die Arbeit nicht ausufern zu lassen und sie zu einem ansprechenden Ende zu führen.

Das Ziel meiner Arbeit war eine möglichst umfassende und kompakte Darstellung der Forschung zum Thema Tonhöhenwahrnehmung mit dem Schwerpunkt auf dem Phänomen Residualtonhören und den aktuellen Erkenntnissen zur neuronalen Autokorrelation. Mein Bestreben war es, die naturwissenschaftlichen Teile in einem musikwissenschaftlichen Kontext verständlich zu präsentieren und eine gute Übersicht zu geben, wie Naturwissenschaften und Musikwissenschaft im Bereich der Tonhöhenwahrnehmungsforschung ineinander greifen. Insofern ist dieses Thema ein Paradebeispiel für Transdisziplinarität in der Forschung, wie sie heute im naturwissenschaftlichen Bereich bereits gang und gäbe ist.

Meine Motivation für dieses Diplomarbeitsthema ist unter anderem auch darin begründet, den naturwissenschaftlich-systematischen Teil der Musikwissenschaft, der ursprüng-

lich die Anfänge der systematischen Beschäftigung mit Musik begründete, zu stärken, um ihn nicht gänzlich zu den Naturwissenschaftlern abwandern zu lassen, sowie eine Brücke zwischen den — einander nicht immer wohlgesinnten — Geistes- und Naturwissenschaften zu schlagen.

Mein Dank gilt Univ.-Prof. Dr. Christoph Reuter, der mir die Anregung gab, mich mit diesem interessanten Thema zu beschäftigen, und mir immer wieder mit Literaturtipps zur Seite stand.

Weiters möchte ich mich bei allen Wegbegleitern bedanken, die bereits seit meiner Kindheit meine Liebe zur Musik geteilt haben; sowie ganz besonders bei denjenigen, die mich bis zum Ende meines Musikwissenschafts-Studiums unterstützt und immer wieder ermuntert haben, diese Arbeit zu Ende zu schreiben. Ihnen sei diese Arbeit gewidmet.

Bettina Wimmer

Wien, am 30. Oktober 2012

Inhaltsverzeichnis

Vorwort	2
Inhaltsverzeichnis	4
1 Einleitung und Begriffsdefinitionen	7
1.1 Der Begriff der Tonhöhe und seine Relevanz	10
1.1.1 Definition des Tonhöhenbegriffs	10
1.1.2 Bedeutung der Tonhöhe für Sprache und Musik	13
1.1.3 Tonhelligkeit und Chroma – Die Zweikomponententheorie	14
1.2 Begriffsdefinitionen	14
2 Aufbau des Ohres und Grundlagen des Hörens	16
2.1 Vom Schall zum Gehirn — Stationen der akustischen Signalverarbeitung	16
2.1.1 Von der Schallquelle zum Ohr	16
2.1.2 Äußeres Ohr und Mittelohr	17
2.1.3 Das Innenohr	17
2.1.4 Die Cochlea	18
2.2 Neurologische Grundlagen der akustischen Signalverarbeitung	20
2.2.1 Aufbau und Kommunikation von Nervenzellen	21
2.2.2 Das Gehirn und der auditorische Cortex	22
2.3 Weiterleitung der akustischen Information vom Trommelfell zum Gehirn .	24
2.4 Wo die Tonhöhenwahrnehmung stattfindet	25
3 Tonhöhenwahrnehmung	27
3.1 Historische Entwicklung der Tonhöhenwahrnehmungsforschung	27
3.1.1 Anfänge der Akustik	27
3.1.2 Der Disput zwischen Ohm und Seebeck über die Tonhöhenwahrnehmung	31
3.1.3 Helmholtz' Beitrag zur Tonhöhenwahrnehmung:	34

3.1.4	Erkenntnisse der Neurologie im 20. Jahrhundert	35
3.1.5	Erkenntnisse der Psychoakustik im 20. und 21. Jahrhundert . . .	37
3.2	Zusammenfassung und aktueller Forschungsstand zur Tonhöhenwahrnehmung	39
3.3	Das Residualtonhören als Phänomen der Tonhöhenwahrnehmung	39
4	Das Residualtonphänomen	42
4.1	Der Residualton – eine Begriffsverwirrung	42
4.2	Die Entdeckung des Residualtonphänomens	42
4.3	Eigenschaften des Residualtons	44
4.3.1	Schoutens Forschungsergebnisse	44
4.3.2	Ritsmas Forschungsergebnisse	46
4.3.3	Wallisers Forschungsergebnisse	48
4.3.4	Weitere Ergebnisse	50
4.4	Residualton versus Kombinationston	52
4.5	Schlussfolgerungen für die Tonhöhenwahrnehmung	53
5	Analyse im Frequenzbereich vs. Analyse im Zeitbereich	54
5.1	Ortstheorie/Analyse im Frequenzbereich	54
5.1.1	Ablauf der Analyse im Frequenzbereich	54
5.1.2	Phänomene, die nicht durch Frequenzbereichsanalyse erklärt werden können	55
5.1.3	Modifikation der Ortstheorie durch Terhardt	56
5.2	Periodizitätstheorie/Analyse im Zeitbereich	58
5.2.1	Ablauf der Analyse im Zeitbereich	58
5.2.2	Phänomene, die nicht durch Zeitbereichsanalyse erklärt werden können	60
5.3	Verknüpfung der beiden Theorien (Frequenz- und Zeitbereichsanalyse) . .	61
6	Mathematische und Neuronale Autokorrelation	63
6.1	Mathematische Autokorrelation	63

6.1.1	Definition der Autokorrelationsfunktion	63
6.1.2	Eigenschaften der Autokorrelationsfunktion	64
6.2	Neuronale Autokorrelation	65
7	Neuronale Modelle der Tonhöhenwahrnehmung	67
7.1	Lickliders <i>Duplex Theory of Pitch Perception</i>	68
7.1.1	Formalisierung des neuronalen Autokorrelationsprozesses nach Lick- lider	68
7.1.2	Biologische Evidenz für ein Autokorrelations-Modell nach Langner	70
7.2	Ein stochastisches Modell der Tonhöhenwahrnehmung von Goldstein . . .	74
7.2.1	Mathematische Formulierung des Goldstein-Modells	75
7.2.2	Der Residualton in Goldsteins Modell	76
7.3	Ein Autokorrelationsmodell von Impulsfolgen	76
7.3.1	Mathematische Formulierung des Autokorrelationsmodells	76
7.3.2	Frickes Ergänzungen zu Unschärfe und Impulsform	78
7.4	Residualton & Neuronale Autokorrelation	82
8	Fazit	84
9	Ausblick	85
10	Literatur	88
11	Abstracts in Deutsch und Englisch	96
11.1	Abstract in Deutsch	96
11.2	Abstract in English	98
12	Lebenslauf der Autorin	99

1 Einleitung und Begriffsdefinitionen

Das Phänomen der Tonhöhenwahrnehmung ist unmittelbar verbunden mit der Frage, was einen Ton überhaupt ausmacht. Obwohl Versuche zur Tonhöhe mit Musikinstrumenten und Saitenlängen bereits vor über 2000 Jahren durchgeführt wurden, war lange Zeit nicht bekannt, wie gewisse wahrgenommene Eigenschaften eines Klanges (bspw. Klangfarbe, Lautstärke und eben auch die Tonhöhe) systematisch beschrieben werden können.

Im 19. Jahrhundert gestatteten neu etablierte mathematische Konzepte wie Integral- und Differentialrechnung und die zunehmende Exaktheit in der Formulierung naturwissenschaftlicher Ideen einen großen Erkenntnisgewinn in vielen Bereichen der Wissenschaft. Es verwundert also nicht, wenn die ersten grundlegenden Resultate auf dem Gebiet der Tonhöhenwahrnehmung von Physikern erlangt wurden, welche sich im 19. Jahrhundert zunehmend intensiv mit dem Gebiet der Akustik auseinanderzusetzen begannen. Von großer Bedeutung war zunächst die Frage nach dem Zusammenhang zwischen den physikalischen Eigenschaften (Frequenz, Amplitude, etc.) und den wahrgenommenen Eigenschaften (Tonhöhe, Lautstärke, Klangfarbe, etc.) eines Tones.

Der Wissenschaftler August Seebeck (1805–1849) schrieb über dieses Forschungsgebiet 1843 in einem Artikel in den *Annalen der Physik und Chemie*:¹

„[...]weil es sich dabei um eine für die Akustik ungemein wichtige Frage handelt, nämlich: was ist als das Wesentliche der Wellen eines Tones anzusehen, sey es daß derselbe allein oder zugleich mit anderen Tönen auftritt?“

Nur wenige Jahre später erschien 1863 mit Helmholtz' Buch *Die Lehre von den Tonempfindungen* ein Standardwerk auf dem Gebiet der Akustik und speziell der Psychoakustik, welches auch heute noch von großer Wichtigkeit ist. In der Einleitung schrieb er über die Einordnung des Themengebiets:²

„Das vorliegende Buch sucht die Grenzgebiete von Wissenschaften zu vereinigen, wel-

¹Seebeck (1843), Seite 449

²von Helmholtz (2003) [1863], Seite 1

che, obgleich durch viele natürliche Beziehungen aufeinander hingewiesen, bisher jedoch ziemlich getrennt nebeneinander gestanden haben, die Grenzgebiete nämlich einerseits der physikalischen und physiologischen Akustik, andererseits der Musikwissenschaft und Ästhetik.“

An der Schwelle zum 20. Jahrhundert kam es somit erstmals zu einer Verknüpfung der formalistischen Sichtweise von Musik (z.B. Gestalt der Schallwelle, Aufbau eines Klanges) mit den subjektiven Empfindungen (z.B. Fragen der Ästhetik), die selbige auslöst.

Im Laufe des 20. Jahrhunderts begannen nach und nach die Physiologen und Mediziner mit akustischen Forschungen. Neue Erkenntnisse aus der Biologie und Neurologie vertieften das Verständnis der Tonhöhenwahrnehmung.

Mit dem Aufkommen der Musikwissenschaft als eigene Disziplin hatte eine weitere Gruppe großes Interesse die Vorgänge beim Hören eines Musikstücks nicht nur außerhalb, sondern auch innerhalb des menschlichen Körpers, zu verstehen. Psychoakustische Forschungsgruppen haben im Laufe der Jahre eine große Zahl an Hörversuchen mit Probanden unternommen, um die Feinheiten und Spezifika des menschlichen Gehörsinns herauszuarbeiten.

Dabei ist insbesondere das Phänomen des Residualtonhörens viele Jahre hindurch beobachtet und analysiert worden. Menschen hören in zahlreichen Situationen Tonhöhen, die physikalisch gesehen nicht vorhanden sind. Während viele andere Eigenschaften des Gehörs bereits hinreichend mit theoretischen Modellen und biologischer Evidenz erklärt werden konnten, gelang es bis vor kurzem nicht, eine zufriedenstellende und physiologisch gut untermauerte Erklärung für das Residualtonhören zu finden.

Mit dem Konzept der neuronalen Autokorrelation und ihrem biologischen Nachweis scheint die Lösung dieser ungeklärten Fragestellung nahe.

Im folgenden Teil dieses Kapitels wird der Begriff der Tonhöhe in der Musikwissenschaft näher beleuchtet, während im dritten Kapitel biologische und neurologische Grundlagen des menschlichen Gehörs präsentiert werden. Im vierten Kapitel kommt es zu einer

ausführlichen Darstellung der historischen Tonhöhenwahrnehmungsforschung, bevor im fünften Kapitel das Residualtonhören und die Forschungsarbeiten dazu genau beschrieben werden. Im sechsten Kapitel passiert ein Vergleich zwischen den Modellansätzen, die eine Frequenzanalyse des Schalls propagieren, im Vergleich zu jenen, die eine Periodizitätsanalyse forcieren. Das siebte Kapitel erklärt die Grundlagen der (neuronalen) Autokorrelation, um das Verständnis der im achten Kapitel vorgestellten Modelle der Tonhöhenwahrnehmung zu erhöhen.

Anschließend folgen Fazit, Ausblick, Abstracts in Deutsch und Englisch sowie der Lebenslauf der Autorin.

Da die Lesbarkeit eines Textes ebenso wie die Gleichstellung der Geschlechter für mich einen hohen Stellenwert hat, verwende ich in weiterer Folge die neutralen Bezeichnungen Wissenschaftler oder Forscher für eine Gruppe von Menschen, die Forschung betreibt, egal welchen Geschlechts diese Menschen sind.

1.1 Der Begriff der Tonhöhe und seine Relevanz

1.1.1 Definition des Tonhöhenbegriffs

Der Begriff der Tonhöhe ist ein Beispiel für eine Bezeichnung, die allgemein bekannt, aber nicht leicht klar fassbar und definierbar ist. Erstaunlicherweise wird in den wenigsten psychoakustischen Büchern die Schwierigkeit der Begriffsbildung erörtert.

Um die Sache zu vereinfachen, wird zunächst der Spezialfall der Sinustöne³ betrachtet. Einem Sinuston kann eine eindeutige Frequenz zugeordnet werden. Die Frequenz in Hertz gibt die Anzahl der Schwingungen pro Minute an. Diese Definition der Tonhöhe von Sinustönen kommt ohne die menschliche Wahrnehmung aus und ist eine objektive Größe. Bei periodischen aus Sinuskomponenten zusammengesetzten Klängen ist es schwieriger, eine Tonhöhe zu definieren. Heutzutage ist bekannt, dass bei derart zusammengesetzten Klängen nicht allein die Frequenz bzw. Periode der Teilschwingungen für die Tonhöhenwahrnehmung relevant ist. A priori ist hier also nicht klar, wie und ob die Tonhöhe eines derartigen Schalls eindeutig festgelegt werden kann.

Noch schwieriger ist die Definition der Tonhöhe für Geräusche oder nicht-periodische Klänge, wie sie bspw. von Glocken erzeugt werden. Hier ist zunächst zu klären, ob es überhaupt eine sinnvolle Tonhöhen-Zuordnung geben kann.

Eine bekannte und weit verbreitete Definition der Tonhöhe (engl. pitch) lieferte die American Standards Association (heute: American National Standards Institute), zuständig für Normungen:⁴

„Pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Pitch depends mainly on the frequency content of the sound stimulus, but it also depends on the sound pressure and waveform of the stimulus.“

Die Tonhöhe wird hier als Wahrnehmungsphänomen definiert, das im Wesentlichen auf den Frequenzanteilen des Schalls beruht. Tramo et al. bezeichnen die obige Definition als

³zur Definition des Begriffs vgl. Abschnitt 1.2

⁴ASA (1960), Seite 44

unvollständig, da sie neue psychoakustische Forschungsergebnisse (vgl. Abschnitt 1.1.3) nicht berücksichtigt.⁵

„The oft-cited, half-century-old definition of pitch by the American National Standards Institute [...] is incomplete, for it implies a single psychological dimension (pitch height). The well-studied perceptual phenomena of octave similarity (pitch chroma) and within-octave pitch-class hierarchies (e.g., tonic-dominant relationships) establish pitch as a multidimensional percept.“

Hier wird die Tonhöhe nicht bloß als eindimensionale Anordnung der Töne von tief bis hoch gesehen, sondern als zweidimensionale Ordnung von pitch chroma und pitch height aufgefasst, die auf der menschlichen Wahrnehmung beruht.

Diesem Umstand wird in der Definition aus *Das neue Lexikon der Musik* Rechnung getragen. Hier wird die Dualität des Tonhöhenbegriffs ausdrücklich betont:⁶

„TONHÖHE zählt neben Tondauer zu den „zentralen“ Eigenschaften eines Tones gegenüber Tonstärke und Klangfarbe als „peripheren“ Eigenschaften (J. Handschin). Akustisch gesehen, ist die T. abhängig von der Anzahl der Schwingungen pro Sekunde (Frequenz) eines elastischen Körpers, z.B. einer Saite: die T. steigt mit wachsender Frequenz. Dabei zeigt T. 2 Komponenten: einerseits die linear sich verändernde Höhe, die an den Eindruck zu- bzw. abnehmender Helligkeit gekoppelt ist; andererseits die Tonigkeit, die solcher Veränderung als zyklisch wiederkehrendes Moment entgegensteht (G. Révész, E.M. von Hornbostel; Gehör).“

*The New Grove*⁷ hingegen gibt folgende Erklärung der Tonhöhe („pitch“), worin sehr stark betont wird, dass Tonhöhe ein subjektives Empfinden des Zuhörers ist, das je nach Situation bei der akustischen Verarbeitung in der Hörbahn unterschiedlich interpretiert wird:

„Pitch. The particular quality of a sound (e.g. an individual music note) that fixes its position in the scale. Certain sounds used in music that occupy no particular scale

⁵Tramo u. a. (2005), Seite 149

⁶lex (1996), Seite 519

⁷Haynes (1995), Seite 779

position, such as those produced by cymbals or the bass drum, can be said to be of indefinite pitch. Pitch is determined by what the ear judges to be the most fundamental wave-frequency of the sound (even when, as for example with difference tones, this is an aural illusion, not actually present in the physical sound wave). Experimental studies, in which listeners have been tested for their perception and memory of pitch differences among sounds with wave-frequencies known to the experimenter, have shown that marked differences of timbre, loudness and musical context affect pitch, albeit in relatively small degree.“

In obiger Definition wird die Erklärung des Tonhöhenbegriffs („pitch“) auf die Klärung der Frage des Begriffes „scale“ (Tonleiter) abgewälzt. Um einen Zirkelschluss zu vermeiden („Tonleiter ist die Anordnung der Töne nach Tonhöhe.“) ist es nötig, einen anderen Weg zu suchen.

Hartmann umgeht die Problematik, die Definition der Tonhöhe in Worte zu fassen, und schlägt eine systematische Vorgehensweise als Definition vor:⁸

„Better than a verbal definition of pitch is an operational definition. We say that a sound has a certain pitch if it can be reliably matched by adjusting a sine wave of arbitrary amplitude. [...] This operational definition leads to a unit for measuring pitch; pitch is measured in Hertz. [...] This operational definition of pitch has several features. First, we are prepared to accept variability. [...] It might also happen that the distribution of pitch matches is bimodal, and in that case we are prepared to say that the pitch is ambiguous or that there are two pitches. [...] Then we conclude that a given stimulus has a certain pitch only if a great majority of listeners make similar matches.“

Diese Definition erscheint sehr praktikabel. Einerseits ermöglicht sie auch zusammengesetzten Klängen und manchen Geräuschen eine Tonhöhe zuzuordnen. Andererseits berücksichtigt sie die Uneindeutigkeit der Tonhöhe in verschiedenen Kontexten. So können Zuhörer dasselbe akustische Signal je nach musikalischer Erfahrung unterschiedlich

⁸Hartmann (1996), Seite 3493

interpretieren. Zum Beispiel gibt es Untersuchungen, dass manche Menschen einen musikalischen Akkord eher holistisch (synthetisch) wahrnehmen und für andere die spektralen Komponenten stärker in den Vordergrund treten.⁹

Gleichzeitig ist durch die individuelle Bewertung, die in die Definition eingeht, stets klar, dass Tonhöhe eine subjektive Wahrnehmung ist, die a priori nicht als objektive Größe angesehen werden kann.

In den meisten Fällen decken sich die individuellen menschlichen Wahrnehmungen jedoch weitgehend. Das freut die Musikwissenschaft: Eine systematische psychoakustische Forschung wäre sonst nicht möglich.

1.1.2 Bedeutung der Tonhöhe für Sprache und Musik

Die Tonhöhe spielt in Alltagssituationen bei der Verwendung der deutschen Sprache nur eine sekundäre Rolle. Üblicherweise dient der Tonhöhenverlauf nur der Unterscheidung zwischen Aussagesatz und Fragesatz („Du kommst?“ vs. „Du kommst.“) bzw. einer damit verbundenen Emotion, sowie zur Wiedererkennung einer menschlichen Stimme. Die semantische Bedeutung eines Lautes wird von der Tonhöhe nicht beeinflusst. In Tonsprachen (z.B. Mandarin, Vietnamesisch) hängt hingegen die Bedeutung eines Lautes von der verwendeten Tonhöhe ab.

Die Tonhöhenwahrnehmung ist also ein vom Menschen ständig verwendeter Prozess, der für die zwischenmenschliche Kommunikation eine große Rolle spielt.

Im musikalischen Zusammenhang ist die Tonhöhe in Form von Melodie und Harmonie von zentraler Bedeutung. Es ist unbestritten, dass die Tonhöheneigenschaft eines musikalischen Klanges entscheidend für dessen Rolle in einem Musikstück ist. Dabei ist jedoch meist nur die relative Tonhöhe entscheidend, also die Beziehung zu den in unmittelbarer zeitlicher Nähe auftretenden Klängen. Hingegen ist die absolute Tonhöhe für Menschen ohne absolutes Gehör von geringer Bedeutung, um eine Melodie oder Harmoniefolge als solche wieder zu erkennen.

⁹Schneider u. Wengenroth (2009)

1.1.3 Tonhelligkeit und Chroma – Die Zweikomponententheorie

Die Zweikomponententheorie der Musikpsychologie knüpft an die bereits erwähnte Problematik der Tonhöhenzuschreibung bei Geräuschen und zusammengesetzten Klängen an.

Sie besagt, dass die Höhe eines Tons durch zwei charakteristische Eigenschaften beschrieben werden kann. Die Tonhelligkeit bestimmt die Zugehörigkeit eines Tones zu einer Oktave. Das Chroma gibt an, welche Position ein Ton innerhalb einer Oktave einnimmt.

In der Zweikomponententheorie wird durch die Einführung des Chroma-Begriffes die besondere Rolle der Oktave in den unterschiedlichsten musikalischen Systemen und Musikkulturen unterstrichen.

Während die Tonhelligkeit eine nach oben offene, lineare Skala ist, bezeichnet im Gegensatz dazu das Chroma eine zyklische Komponente. Geräuschen kann kein Chroma zugeordnet werden, allerdings kann auch ihnen eine Tonhelligkeit zugeschrieben werden.

Diese Arbeit beschäftigt sich in weiterer Folge mit der linearen Tonhöhe. Eine ausführliche Behandlung der Zweikomponententheorie findet sich bspw. bei Ebeling.¹⁰

1.2 Begriffsdefinitionen

In dieser Arbeit werden die Begriffe Schallwelle, Schallsignal, Schallereignis, akustischer Reiz, akustisches Signal, etc. im Wesentlichen äquivalent für einen mit den Ohren wahrnehmbaren Sinneseindruck verwendet.

Die Bezeichnung Ton oder Klang verwende ich für ein allgemeines akustisches Signal bzw. eine nicht näher spezifizierte periodische Schallwelle. Im Gegensatz dazu steht Sinuston oder einfacher Ton für ein sinusförmiges Schallsignal. Mit komplexer Klang, musikalischer Klang bzw. zusammengesetzter Klang ist grundsätzlich ein allgemeines, aus sinusförmigen Komponenten zusammengesetztes Schallsignal gemeint. Das bedeutet, dass dieser Begriff großteils ein beliebiges periodisches Schallereignis beschreibt und

¹⁰Ebeling (1999)

somit generell auch als Spezialfall rein sinusförmige Schallsignale (Sinustöne) umfasst. In manchen Situationen wird der Begriff allerdings für eine Gegenüberstellung sinusförmiger Schallsignale und nicht-sinusförmiger, aber periodischer Schallsignale verwendet. Die jeweilige Bedeutung wird aus dem Kontext stets klar erkennbar sein. Ein harmonischer Klang bezeichnet speziell einen Grundton mit den dazugehörigen Obertönen.

2 Aufbau des Ohres und Grundlagen des Hörens

In diesem Kapitel werden die biologischen und neurologischen Grundlagen des menschlichen Gehörsinns zusammengefasst, mit besonderem Augenmerk auf jene Bereiche, die für die Tonhöhenwahrnehmung und das Residualtonhören relevant sind.

2.1 Vom Schall zum Gehirn – Stationen der akustischen Signalverarbeitung

2.1.1 Von der Schallquelle zum Ohr¹¹

Ein musikalischer Klang oder ein Geräusch pflanzt sich durch Schallwellen in einem elastischen Medium fort. Diese Schallwellen entsprechen Druckschwankungen im Übertragungsmedium und sind in gasförmigen oder flüssigen Medien stets Longitudinalwellen. Das bedeutet, dass die Auslenkung in Fortpflanzungsrichtung erfolgt.¹²

Für jedes Medium gilt eine charakteristische Ausbreitungsgeschwindigkeit des Schalls, die Schallgeschwindigkeit c , die von Faktoren wie Temperatur, Dichte oder Aggregatzustand abhängt.¹³

Da für den menschlichen Gehörsinn primär das Medium Luft relevant ist, wird in weiterer Folge ausschließlich diese Situation beleuchtet.

Luftmoleküle sind in ständiger, rein zufälliger Bewegung. Breitet sich von einer punktförmigen Quelle eine Schallwelle kugelförmig aus, so werden die Luftmoleküle für kurze Zeit zu einer Oszillation radial zur Ausbreitungsrichtung angeregt. Es kommt zu Dichte- und Druckschwankungen in der Luft. Der Schalldruck wird mit einer Geschwindigkeit von ca. 344 m/s an benachbarte Moleküle weitergegeben. Der Abstand zwischen zwei Druckmaxima ist die Wellenlänge λ . Die Anzahl der Druckmaxima pro Sekunde wird die Frequenz f genannt.¹⁴

¹¹Gelfand (2009), Seite 8–20

¹²Gelfand (2009), Seite 10–11

¹³Rieländer (1982), Seite 313

¹⁴Gelfand (2009), Seite 16

Es gilt für jedes Medium:¹⁵

$$\lambda \cdot f = c$$

Die gesamte Schallinformation wird rein durch diese minimalen Luftdruckschwankungen übermittelt.

Durch die Ohrmuschel und den Gehörgang gelangen sie in das Innenohr und treffen dort auf das Trommelfell, wo der Prozess des Hörens, also die Analyse akustischer Signale beginnt.

2.1.2 Äußeres Ohr und Mittelohr¹⁶

Der einzige von außen sichtbare Teil des Ohres ist die Ohrmuschel (Auricula), die hauptsächlich für das Richtungshören und für die Bündelung des Schalls verantwortlich ist. Unmittelbar daran schließt der mit Luft gefüllte Gehörgang an, der am inneren Ende von der dünnen Haut des Trommelfells begrenzt wird. Das Trommelfell ist eine Membran, die durch den Schall in Bewegung versetzt wird. Es wandelt die Luftdruckschwankungen in eine mechanische Schwingung im Mittelohr um, die wiederum die Gehörknöchelchenkette - bestehend aus Hammer (Malleus), Amboss (Incus) und Steigbügel (Stapes) - in der Paukenhöhle anregt. Die dadurch entstehende Hebelwirkung verstärkt die Luftdruckschwankungen, um diese auf die Lymphe des Innenohres zu übertragen. Dies ist notwendig, da die Lymphe ein dichteres Medium als die Luft im Mittelohr ist und somit höhere Energien zur Schallausbreitung erforderlich sind.

Die Eustachische Röhre ermöglicht bei Bewegung des Trommelfells den Druckausgleich zwischen Mittelohr und Rachenraum.

2.1.3 Das Innenohr¹⁷

Das mit Perilymphe gefüllte Innenohr wird wegen seiner vielen Verzweigungen auch „Labyrinth“ genannt und ist in einem der härtesten menschlichen Knochen, dem Felsenbein,

¹⁵Gelfand (2009), Seite 16

¹⁶Fastl u. Zwicker (2007), Seite 23–25; Hesse (2003b), Seite 65; Yost (1994), Seite 61ff

¹⁷Fastl u. Zwicker (2007), Seite 25; Hesse (2003b), Seite 65; Yost (1994), Seite 75ff

eingebettet. Das Innenohr ist einerseits für den Gleichgewichtssinn (Vestibularapparat) und andererseits für das Hören (Cochlea) zuständig. Über zwei Öffnungen ist es mit dem Mittelohr verbunden. Eine davon ist das ovale Fenster, das von der Fußplatte des Steigbügelknochens verschlossen wird. Die zweite Öffnung, das runde Fenster, wird von einer dünnen Membran, dem sekundären Trommelfell, verschlossen und dient der Abdämpfung der Schallschwingungen in der Cochlea.

2.1.4 Die Cochlea¹⁸

Die Cochlea hat eine aus zweieinhalb Windungen bestehende Schneckenform, die sich um die Spindel (Modiolus) windet. Ihr Durchmesser verringert sich kontinuierlich in Richtung Spitze (Apex). Sie besteht aus zwei durch eine Trennwand separierten Kanälen, die beide mit Perilymphe gefüllt sind. Über das Helicotrauma, eine winzige Öffnung an der Spitze, sind diese miteinander verbunden.

Der obere Kanal der Schnecke mündet am basalen Ende in den Vorhof des Mittelohres, wo der Steigbügel befestigt ist. Dieser ist für die Übertragung der Schallschwingungen an die Perilymphe verantwortlich.

Der untere Kanal (Scala tympani) beginnt im bereits erwähnten runden Fenster, wo eine Membran die Abgrenzung zur Paukenhöhle sicherstellt.

Die beiden Kanäle werden auf der gesamten Länge der Cochlea durch einen mit Endolymphe gefüllten häutigen Schlauch getrennt. Die Membran an seiner Oberseite heißt Reissnersche Membran, die an der Unterseite Basilarmembran.

Die Basilarmembran: Während sich die knöcherne Schneckenform der Cochlea zur Apex hin zuspitzt, verbreitert sich die Basilarmembran in diese Richtung kontinuierlich.¹⁹

Der Physiologe Georg von Békésy (1899–1972) beschäftigte sich intensiv mit der Vermessung der Basilarmembran und fand dabei heraus, dass diese am basalen Ende wesentlich

¹⁸Fastl u. Zwicker (2007), Seite 25; Hesse (2003b), Seite 66; Yost (1994), Seite 75ff

¹⁹Yost (1994), Seite 79

steifer als an der Spitze ist. Diese beiden Eigenschaften der Basilarmembran sind entscheidend für die Tonhöhenwahrnehmung, da sie beeinflussen, wo in der Cochlea sich das Maximum der durch die Schallwellen verursachten Wanderwellen ausbildet.²⁰

Auf der Basilarmembran sitzt das eigentliche Sinnesorgan, das Cortische Organ, benannt nach seinem Entdecker Alfonso Corti (1822–1876).²¹

Das Cortische Organ: Das Cortische Organ ist ein Stützzellengerüst, in dem sich die Haarzellen befinden. Deren feine Härchen (Stereocilia) sind für die Umwandlung eines mechanischen Reizes in Nervenimpulse verantwortlich.²²

Es gibt ca. 12000–15000 äußere Haarzellen²³, die in drei Reihen an der Walloberseite des Cortischen Organs liegen, und ca. 3500 innere Haarzellen. Auf jeder inneren Haarzelle sind ca. 40 Härchen, während sich auf einer äußeren Haarzelle ca. 150 davon befinden.²⁴

Heutzutage gilt als erwiesen, dass die äußeren Haarzellen eine wichtige Rolle als Verstärker bei der Frequenzkomponentenanalyse spielen.²⁵

„Erst in jüngster Zeit ist klar geworden, dass es sich dabei [Anm.: gemeint ist der aktive Verstärkermechanismus des Innenohres] um die in 3 Reihen angeordneten 12000 äußeren Haarzellen handelt (Hudspeth 1985). Ihre Funktion war noch vor etwa 20 Jahren ein Rätsel - sind es doch fast ausschließlich (zu 95%) die nur 3500 inneren Haarzellen, welche die Hörinformation an das Gehirn weiterleiten. Die durch die Schallwellen ausgelöste Depolarisation der äußeren Haarzellen veranlasst diese, sich aktiv im selben Rhythmus zu kontrahieren. Dies verstärkt nicht nur die Amplitude, sondern auch die Frequenzabstimmung der Basilarmembran.“

Evolutionsbiologisch gesehen haben höher entwickelte Wesen eine längere Basilarmembran und mehr Härchen. Békésy untersuchte die Auswirkungen dieser biologischen Un-

²⁰von Békésy (1970), Seite 158

²¹Hesse (2003b), Seite 67

²²Hesse (2003b), Seite 67

²³Langner (2007), Seite 12 bzw. Hesse (2003b), Seite 67

²⁴Yost (1994), Seite 83

²⁵Langner (2007), Seite 12

terschiede auf die Tonhöhenwahrnehmung:²⁶

„Daher muß eine Beziehung zwischen der Länge der Basilarmembran und dem verbesserten Gehörsinn der höheren Tiere vorhanden sein. Da die absolute Empfindlichkeit im Sinne des Minimums der wahrnehmbaren Druckamplitude bei allen Tieren ungefähr gleich ist, muß die größere Länge der Basilarmembran mit der besseren Tonhöhenunterscheidung bei höheren Tieren in Verbindung stehen, einem Unterscheidungsvermögen, das eine große biologische Bedeutung im Verstehen von Sprechgeräuschen annimmt.“

Die Spindel und der Hörnerv: Die von der Cochlea umwickelte Spindel beinhaltet mit dem Ganglion spirale eine große Anzahl bipolarer Nervenzellen, d.h. Neuronen mit zwei Fortsätzen. Einer der beiden Fortsätze führt zum Cortischen Organ und erhält dort Reize von der Basis der äußeren und inneren Haarzellen.²⁷

Ein durch ein akustisches Ereignis ausgelöster Reiz wird auf diese Weise vom Innenohr an das Zentralnervensystem (ZNS) übertragen. Die zu einem Nervenfaserbündel vereinigten Nerven formen den Hörnerv, der wiederum gemeinsam mit dem Gleichgewichtsnerv den Nervus vestibulocochlearis (VIII. Hirnnerv) bildet. Dieser führt über die Schneckenkerne (Nuclei cochleares), die oberen Oliven (Nuclei olivares), wo erstmals Signale des rechten und linken Ohres zusammentreffen, und die Colliculi inferiores – alle diese Stationen sind Teile des Hirnstammes – zu den medialen Kniehöckern (Corpus geniculatum mediale) im Metathalamus (Teil des Zwischenhirns) und weiter zum auditorischen Cortex.²⁸

2.2 Neurologische Grundlagen der akustischen Signalverarbeitung

Der Gehörsinn ist Teil des senso-motorischen Nervensystems. An der Verarbeitung der akustischen Information sind dabei das periphere Nervensystem (PNS), zu dem u.a. der

²⁶von Békésy (1970), Seite 156–157

²⁷Hesse (2003b), Seite 67

²⁸Yost (1994), Seite 228ff

Hörnerv gehört, und das Zentralnervensystem (ZNS), das sich aus Rückenmark und Gehirn zusammensetzt, beteiligt. Bei den Nervenbahnen, die Informationen vom Körper in Richtung Gehirn weiterleiten (u.a. Hörnerv) spricht man von afferenten Nervenbahnen.²⁹

2.2.1 Aufbau und Kommunikation von Nervenzellen

Eine Nervenzelle (Neuron) dient der Reizweiterleitung und besteht aus dem Zellkörper sowie den davon abgehenden Dendriten und dem Axon. Diese Teile ermöglichen den Informationsaustausch mit anderen Neuronen. Während die Dendriten kurze Fortsätze sind, deren Rezeptoren dem Empfang von Informationen dienen, wird über das längliche Axon das elektrische Aktionspotential weitergeleitet.³⁰

Die Schnittstellen zwischen zwei Neuronen bilden Synapsen. Dort sorgen Botenstoffe, sogenannte Neurotransmitter, für die chemische Übertragung des Aktionspotentials (1 mV, 1ms)³¹ und somit für die Reizweiterleitung zwischen den Nervenzellen.

Ein Neuron kann nur zwei Zustände aufweisen: feuernd oder nicht feuernd. Das Aktionspotential ist stets gleich groß. Die Kommunikation zwischen Nervenzellen kann daher binär beschrieben werden: 1 = feuernd, 0 = nicht feuernd. Zusätzliche Informationen (z.B. die Intensität eines Reizes) müssen über die Frequenz codiert werden. Nach Auslösung des Aktionspotentials dauert es einen kurzen Zeitraum, bis die Zelle wieder in einem feuerbereiten Zustand ist. Dieser Zeitraum wird Refraktärzeit genannt. Um die Feueraten von Neuronen vergleichen zu können, verwendet man üblicherweise ISI (interspike interval)-Histogramme, die die zeitlichen Abstände zwischen zwei aufeinanderfolgenden Neuronenentladungen nach Häufigkeit darstellen. Im Gegensatz dazu beruht ein all-order-ISI-Histogramm auf den Abständen zwischen allen möglichen Kombinationen von zwei Neuronenentladungen.³²

Gewöhnlich kommt es erst zur Auslösung eines Aktionspotentials, wenn eine Nervenzelle von mehreren anderen angeregt wird.³³

²⁹Hesse (2003b), Seite 14–15

³⁰Hesse (2003b), Seite 17

³¹Langner (2007), Seite 12

³²Tramo u. a. (2001), Seite 191

³³Fricke (2011), Seite 104

„Generally the threshold of excitement for discharging an impulse in a nerve cell is not reached by a single impulse from a pre-synaptic cell. It takes several input impulses within the integration period to raise the electric potential of a cell to a value that is required for triggering. The integration period relates to the refractory period, which in turn can be described by the decay time of the action potential. The decay time must be taken as approximately 5 msec. Therefore, the result of the summation of the input potential is higher, the smaller the distance between the input impulses. An integration period of 1 msec can be considered as guide value for succesful summation.“

Neuronen ermöglichen auf diese Weise die gezielte Informationsweiterleitung im menschlichen Körper. Bei der Verarbeitung akustischer Reize arbeiten verschiedene Typen von Nervenzellen zusammen (vgl. Abschnitt 7.1.2).

2.2.2 Das Gehirn und der auditorische Cortex

Das menschliche Gehirn besteht aus ca. 10^{11} Neuronen, wobei jedes von ca. 20000 anderen Neuronen Signale empfängt und an ebenso viele weiterleitet.³⁴ Es besteht aus dem Großhirn (Telencephalon), dem Kleinhirn (Cerebellum), dem Zwischenhirn (Diencephalon) und dem Hirnstamm (Truncus cerebri). Das Großhirn umfasst die rechte und die linke Hemisphäre, die durch mehrere Balken miteinander verbunden sind, sowie die darunter liegenden Kerne.³⁵

Der auditorische Cortex ist Teil der Großhirnrinde und die Endstation der Hörbahn. Er dient der Verarbeitung und Analyse akustischer Signale. Beim Menschen liegt der auditorische Cortex am Temporallappen; großteils in der Sylvischen Furche, die sich seitlich des Kopfes in beiden Hemisphären befindet.³⁶

Im Gegensatz zur Benennung anderer Teile des Gehirns umfasst der Begriff „auditorischer Cortex“ somit kein biologisch abgegrenztes Gebiet, sondern Bereiche aus unter-

³⁴Fricke (2011), Seite 104

³⁵Benninghoff u. Drenckhahn (2004), Seite 233ff

³⁶Benninghoff u. Drenckhahn (2004), Seite 740f

schiedlichen Gehirnregionen, wie Tramo et al. ausführen:³⁷

„The term auditory cortex does not refer to a specific brain structure(s) that can be identified directly via static magnetic resonance images, pathological specimens, or histological sections. By definition, auditory cortex refers to gray matter of the cerebral cortex whose neurons respond to auditory stimuli but not to visual, tactile, or other sensory stimuli.“

Der Begriff auditorischer Cortex beschreibt somit alle am Hören beteiligten Gehirnstrukturen. Bei Untersuchungen wurde beobachtet, dass an der Tonhöhenwahrnehmung hauptsächlich die rechte Hälfte beteiligt ist:³⁸

„Die linke Hemisphäre arbeitet sequenziell — also Schritt für Schritt — und ist daher bei der Verarbeitung der aufeinanderfolgenden Worte der Sprache oder bei algebraischen Denkopoperationen dominant, ebenso beim Erfassen des musikalischen Rhythmus, während die entsprechenden Zentren der rechten Seite holistisch, ganzheitlich arbeiten. Sie treten beim Erkennen klanglicher Aspekte in den Vordergrund und sind bei der Verarbeitung von musikalisch-klanglicher Primärinformation aber auch beim Verstehen des Sprachklanges, der Betonung tätig.“

Der auditorische Cortex weist in mehreren Bereichen eine räumliche Anordnung der Tonhöhe (Tonotopie) auf.³⁹ So befinden sich im primären auditorischen Rindenfeld (Area 41) und im sekundären auditorischen Rindenfeld (Area 42)⁴⁰ gegenläufige tonotope Abbildungen. Während ersteres für die Verarbeitung von Information aus der Cochlea verantwortlich ist, finden sich in den Sekundärfeldern Gedächtnisbilder von Sinneseindrücken, sodass in diesem Bereich eintreffende Reize in Sinneseindrücke umgewandelt werden können.⁴¹ In weiterer Folge übernimmt das limbische System die Bewertung des Gehörten und eine etwaige Reaktion darauf (z.B. das Auslösen einer Emotion).⁴²

³⁷Tramo u. a. (2005), Seite 152

³⁸Hesse (2003b) Seite 21

³⁹Benninghoff u. Drenckhahn (2004), Seite 740–741

⁴⁰Die Bezeichnungen *Area 41* und *Area 42* entsprechen der Brodmann-Terminologie zur Bezeichnung von Arealen der Großhirnrinde (Cortex cerebri).

⁴¹Benninghoff u. Drenckhahn (2004), Seite 236–237

⁴²Hesse (2003b), Seite 22

2.3 Weiterleitung der akustischen Information vom Trommelfell zum Gehirn

Wenn eine Schallwelle auf das Trommelfell stößt, wird durch die Vibrationen mechanisch eine Bewegung des Steigbügels ausgelöst. Durch diese Schwingung entsteht in der Cochlea eine Wanderwelle, die sich an der Basilarmembran ausbreitet.⁴³ Je tiefer die zur Wanderwelle zugehörige Frequenz, desto langsamer und weiter läuft die Welle in Richtung Schnecken spitze. Die Wanderwellen erreichen hierbei eine Geschwindigkeit von ca. 1 bis 3m/sec.⁴⁴

Dabei entsteht an einem bestimmten Ort ein Amplitudenmaximum. Kurz darauf stoppen Reibungskräfte das weitere Fortpflanzen der Wanderwelle.

Greenwood⁴⁵ verwendete Daten aus Untersuchungen der Cochlea bei Säugetieren (u.a. Katzen, Meerschweinchen, aber auch Menschen) und beschrieb damit die Frequenzverteilung entlang der Basilarmembran. Beim Menschen ergab sich dabei die bekannte örtliche logarithmische Frequenzanordnung ab etwa 100 Hz und darüber. Darunter nimmt die Frequenz langsamer zu, die Abstände sind dort somit größer als erwartet.

Die Cochlea arbeitet technisch gesehen wie ein seriell geschaltetes Tiefpass-Filter-System, bei dem die obere Frequenzgrenze ständig abnimmt. Eine moderne Interpretation der Basilarmembran wäre somit die Beschreibung als Filterbank, wie Fricke es formulierte:⁴⁶ „Die Basilarmembran wird üblicherweise als Filterbank mit ziemlich breiten Filtern vorgestellt. 24 von ihnen passen nebeneinander in den Hörbereich von unten bis oben, wenn man sie mit festen Grenzen nebeneinander legen würde. [...] Als Richtwert geht man von einer Intervallbreite von $\frac{f_{go}}{f_{gu}} = 1,22$ aus [Anm.: f_{go} = Frequenz der oberen, f_{gu} = Frequenz der unteren Intervallgrenze], was etwa einer neutralen Terz entspricht. Unterhalb von 500 Hz haben die Frequenzgruppen eine Breite von etwa 100 Hz. [...] Innerhalb dieser Frequenzgruppen [...] findet eine gegenseitige Beeinflussung der Schwingungen statt.

⁴³Langner (2007), Seite 11

⁴⁴Fricke (2011), Seite 107

⁴⁵Greenwood (1990)

⁴⁶Fricke (2009), Seite 103

[...] Die Positionen dieser Filterbereiche sind in Wirklichkeit aber variabel, sie passen sich dem jeweils eintreffenden Schall an und gruppieren sich um die analysierten Schallkomponenten ständig neu.“

Die Bewegungsenergie der Schallwelle wird anschließend von den inneren Haarzellen in einen Nervenimpuls umgewandelt.

An der Stelle des Amplitudenmaximums verursacht die Bewegung der Basilarmembran, dass die Stereozilien der Haarzellen verbogen werden. Dadurch kommt es zur Aufnahme von positiv geladenen Kalium-Ionen,⁴⁷ die eine Depolarisation der Haarzelle auslösen. Je höher die Amplitude des Schallsignals, desto stärker ist die Verbiegung. Auf diese Weise kann die Amplitude in einer Abfolge von Nervenimpulsen codiert werden. Anschließend sorgen Transmitterstoffe für die Informationsweiterleitung an die Hörnervenzellen, die wiederum mittels Aktionspotential für die Weiterleitung zur Großhirnrinde sorgen.

Die in der Basilarmembran gewonnene Tonotopie bleibt dabei auf dem gesamten Verarbeitungsweg erhalten.⁴⁸

„Ein komplexes Signal wird auf diese Weise nach seinen Frequenzen zerlegt und als Erregungsmuster entlang der Cochlea abgebildet. Diese räumliche Repräsentation von akustischen Signalen (Tonotopie) wird über den Hörnerven auf die nächste Hörbahnstation, den Nucleus cochlearis, und sukzessive auf alle weiteren Hörbahnstationen bis hinauf zum Hörcortex beibehalten.“

2.4 Wo die Tonhöhenwahrnehmung stattfindet

Spätestens seit Mitte des 20. Jahrhunderts gilt als gesichert, dass sowohl die Cochlea als auch die Basilarmembran eine zentrale Rolle in der Tonhöhenwahrnehmung spielen. Die Forschungsarbeiten von Georg von Békésy zur Steifigkeit und Wellenausbreitung auf der Basilarmembran waren von großer Bedeutung für diese Erkenntnis. Er erhielt dafür 1961 den *Nobelpreis für Physiologie oder Medizin*.

⁴⁷Langner (2007), Seite 12

⁴⁸Langner (2007), Seite 12

Neuere Forschungsergebnisse lassen darauf schließen, dass die akustische Information zusätzlich einer zeitlichen Analyse auf der neuronalen Ebene unterzogen wird (vgl. Abschnitt 5.2). Der exakte Ort und Ablauf der Tonhöhenwahrnehmung im Gehirn ist noch nicht vollständig geklärt, es gibt aber Hinweise darauf, dass speziell die Gyri temporales transversi eine wichtige Rolle spielen, wie Bizley und Walker festhalten:⁴⁹

„The above investigations provide compelling evidence that the lateral HG [Anm.: Gyri temporales transversi sind Teile des auditorischen Cortex] plays a key role in pitch perception, but this does not mean that other cortical regions do not contribute to pitch perception. Further studies suggest that neural representations of pitch may be distributed across multiple regions of the auditory cortex, with different areas contributing to different types of pitch judgments.“

⁴⁹Bizley u. Walker (2010), Seite 6

3 Tonhöhenwahrnehmung

3.1 Historische Entwicklung der Tonhöhenwahrnehmungsforschung

3.1.1 Anfänge der Akustik

Dieses Kapitel beschreibt die historische und aktuelle Entwicklung in der akustischen Forschung zum Thema Tonhöhenwahrnehmung.

Die ersten Erkenntnisse über die Saitenlängen eines Instruments und die damit verbundene erklingende Tonhöhe fanden sich bereits auf Keilschrifttexten der Sumerer⁵⁰ und später in der Antike bei den Griechen. Das Monochord diente Gelehrten wie Pythagoras oder Euklid als Forschungsgegenstand. So war insbesondere die Tatsache, dass eine halbierte Saite genau eine Oktav höher klingt, bereits vor Christi Geburt bekannt, ebenso wie die Tatsache, dass raschere Schwingungen zu einer höheren Tonhöhe führen.⁵¹

Doch erst 1500 Jahre später waren mit dem Aufkommen der Naturwissenschaften, speziell der Physik, die Möglichkeiten gegeben, Töne und Klänge zu charakterisieren und mathematisch zu beschreiben. Mit der zunehmenden Mathematisierung der Physik und der Naturwissenschaften allgemein, hielten Formeln auch im Forschungsgebiet der Akustik Einzug.⁵²

Man begann Gleichungen für die Schwingung einer Saite, eines Trommelfells und der Luftmoleküle bei einer Anregung durch Schallwellen aufzustellen. Gleichzeitig tauchte die Frage auf, wie die Eigenschaften des Schalls (Lautstärke, Tonhöhe, etc.) in diese Formeln einfließen.

⁵⁰Hesse (2003b), Seite 112

⁵¹Dostrovsky u. Cannon (1987), Seite 11

⁵²Man denke etwa an die Verwendung von Integral- und Differentialrechnung, die Ende des 17. Jahrhunderts von Newton und Leibniz formuliert wurden und die sich im 18. Jahrhundert als nützliches Werkzeug zur Beschreibung von dynamischen Vorgängen in der Physik (z.B. Geschwindigkeit, Beschleunigung,...) erwiesen.

Der italienische Physiker Galileo Galilei (1564–1642) beschäftigte sich unter anderem intensiv mit Saiten und beschrieb die Abhängigkeit der Tonhöhe von deren Beschaffenheit.⁵³

„He [Galileo] then proceeds to explain the relation of pitch to frequency, consonance and dissonance and the frequency ratios corresponding to musical intervals, vibratory resonance, sympathetic vibrations and the quantitative dependence of the frequency of vibration of a string on its length, diameter, density and tension. Although none of these topics was new, a few of his explanations were, and the compact presentation of the whole subject in the form of objective conclusions demonstrably based on experiment was a highly significant contribution.“

Weniger bekannt ist der niederländische Philosoph Isaac Beeckman (1588–1637), der mit dem französischen Wissenschaftler Descartes (1596–1650) in regem Austausch über akustische Fragestellungen stand und bereits 1614 den Zusammenhang zwischen Frequenz und Tonhöhe schriftlich formuliert hatte.⁵⁴

Der französische Gelehrte Marin Mersenne (1588–1648) beobachtete wenige Jahre später bei verschiedenen Instrumenten (u.a. Saiteninstrumente, Trompeten, Glocken) die sogenannten „Nebentöne“, die heute unter dem Namen „Obertöne“ bekannt sind.⁵⁵ Wichtige wissenschaftliche Beiträge zu Obertönen lieferte auch der Franzose Joseph Sauveur (1653–1716). Er beschäftigte sich mit dem Orgelbau und erkannte, dass ein musikalischer Klang stets aus dem Grundton und den harmonischen Obertönen zusammengesetzt ist.⁵⁶ Mit Hilfe der Schwebung zweier Orgelpfeifen versucht er sich an Absolutfrequenzmessungen.⁵⁷ Weiters verdanken wir Sauveur die Einführung der Begriffe Akustik, harmonischer Klang, Schwingungsknoten und -bauch, sowie die Idee, dass Klangfarbe und Obertöne zusammenhängen.⁵⁸

Ungefähr zu selben Zeit gelang es dem Physiker und Mathematiker Isaac Newton (1642–

⁵³Hunt (1978), Seite 80 und 81

⁵⁴Dostrovsky u. Cannon (1987), Seite 21

⁵⁵Dostrovsky u. Cannon (1987), Seite 39

⁵⁶Hesse (2003b), Seite 113

⁵⁷Dostrovsky u. Cannon (1987), Seite 35

⁵⁸Dostrovsky u. Cannon (1987), Seite 46–47

1727) die Schallausbreitung im Raum zu beschreiben.⁵⁹ Für seine Berechnungen verwendete er eine Schallgeschwindigkeit in der Luft von ca. 346,5 m/s, die 1708 der englische Theologe William Derham (1657–1735) bestimmt hatte. Dieses Ergebnis entspricht schon beinahe dem heutigen Wert von ca. 344 m/s bei Umgebungstemperatur auf der Erde. Und auch die Gleichung $\lambda f = c$ (Produkt aus Wellenlänge und Frequenz ist stets konstant und entspricht der Schallgeschwindigkeit) war Newton bereits bekannt.⁶⁰

Das 18. Jahrhundert war stark geprägt von der Theorie der Differentialgleichungen und gleichzeitig der Beginn der Fourier-Analyse („harmonische Analysis“, Theorie der Fourierreihen, vgl. Formel (1)).

Die führenden Naturwissenschaftler (u.a. Brook Taylor (1685–1731), Johann (1667–1748) und Daniel Bernoulli (1700–1782), Johann Albrecht Euler (1734–1800), Jean-Baptiste le Rond D’Alembert (1717–1783)) entwickelten Konzepte zur Beschreibung schwingender Saiten, schwingender Luftsäulen, schwingender Stäbe und hängender Ketten.⁶¹ Schließlich stellte 1746 D’Alembert die berühmte Wellengleichung auf.⁶²

Systematische psychoakustische Untersuchungen des menschlichen Gehörs und somit auch der Tonhöhenwahrnehmung fanden erst im 19. Jahrhundert statt. Mit der vom französischen Erfinder Charles Cagniard de Latour (1777–1859) 1819 erfundenen Sirene war es erstmals möglich, Sinustöne mit einer exakten Frequenz zu erzeugen und diese beliebig zu reproduzieren, wie Turner festhielt:⁶³

„Before [...], tones of a specific frequency or frequency interval could be produced only by sounding musical notes or chords on standard instruments, a technique both subjective and imprecise. Cagnard’s [sic] siren, especially as improved versions of it appeared, allowed tones of precise frequency to be readily generated mechanically. Experimentalists all over Europe used the siren, but by the early 1840s it had attained a special importance in Germany [...].“

⁵⁹Dostrovsky u. Cannon (1987), Seite 65

⁶⁰Dostrovsky u. Cannon (1987), Seite 42

⁶¹Dostrovsky u. Cannon (1987), Seite 50ff

⁶²Dostrovsky u. Cannon (1987), Seite 66

⁶³Turner (1977), Seite 2

Dadurch ergab sich ein breites Betätigungsfeld für Forscher wie August Seebeck, Hermann von Helmholtz (1821–1894) oder Arthur König (1856–1901)⁶⁴, die mit Hilfe von Lochsirenen zahlreiche Hörexperimente durchführten. Die Erzeugung eines Tons erfolgt dabei durch einen konstanten Luftstrom, welcher auf eine rotierende Lochscheibe bläst. Durch die Abstände der Löcher auf der Scheibe ergibt sich die Frequenz und somit die Tonhöhe.

Die meisten Forscher beschäftigten sich zu dieser Zeit mit einer Vielzahl naturwissenschaftlicher Fragestellungen und nicht nur mit rein akustischen. So hatte etwa der deutsche Physiker Georg Simon Ohm (1789–1854) auch wichtige Beiträge zur Elektrizitätslehre geliefert⁶⁵ und August Seebeck interessierte sich ebenso für optische Fragestellungen.

Während viele mathematische Begriffe für eine konkrete physikalische Fragestellung entwickelt wurden, gelang dem französischen Mathematiker Jean Baptiste Joseph Fourier (1768–1830) das Kunststück, dass ein von ihm entdeckter rein mathematischer Satz später als Grundlage einer akustischen Theorie diente.

Das Fourier-Theorem besagt:⁶⁶

Jede Funktion f mit Periode $2T$ kann in eine Fourierreihe entwickelt werden.

Die Darstellung sieht wie folgt aus:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[a_k \cos \frac{\pi k x}{T} + b_k \sin \frac{\pi k x}{T} \right] \quad (1)$$

$$a_k = \frac{1}{T} \int_{-T}^T f(x) \cos \frac{\pi k x}{T} dx \quad k = 0, 1, 2, \dots \quad (2)$$

$$b_k = \frac{1}{T} \int_{-T}^T f(x) \sin \frac{\pi k x}{T} dx \quad k = 1, 2, \dots \quad (3)$$

Mit der Gleichung $\sin(x) = \cos(x + \frac{\pi}{2})$ kann die Reihe auch rein mit Sinus- bzw. Cosinustermen geschrieben werden.

⁶⁴Turner (1977), Seite 2

⁶⁵vgl. die nach ihm benannte Einheit *Ohm* für den elektrischen Widerstand.

⁶⁶Endl u. Luh (2008), Seite 113ff

Jede periodische Funktion ist somit als (möglicherweise unendliche) Summe von Sinus- bzw. Cosinusfunktionen darstellbar. Zur Bedeutung dieser Erkenntnis für die Musikwissenschaft, vgl. Barkowsky.⁶⁷

Diese Idee inspirierte Forscher im Bereich der Akustik, eine beliebige periodische Schallwelle als Zusammensetzung von Sinuskomponenten zu interpretieren. Für diese heute anerkannte Theorie wurde der Physiker Georg Ohm zunächst noch heftig von seinem Kollegen August Seebeck kritisiert.

3.1.2 Der Disput zwischen Ohm und Seebeck über die Tonhöhenwahrnehmung

1841 veröffentlichte Seebeck einen Artikel über die Resultate seiner Forschungen an der Lochsirene in den *Annalen der Physik und Chemie*.

Es war ihm gelungen, mittels zweier Lochsirenen die Effekte Auslöschung (durch Phasenverschiebung) und Lautstärkenverdopplung (durch gleiche Phase) bei einem Sinuston zu erzeugen.⁶⁸

In einem weiteren Versuch verwendete er eine Scheibe mit zwei Lochreihen, wobei die eine Lochreihe exakt doppelt so viele Löcher wie die andere aufwies. Er blies nun zunächst beide Reihen von derselben Seite an und erzeugte so einen Sinuston und die Oktave dazu. Wenn allerdings die beiden Reihen phasensynchron von verschiedenen Seiten angeblasen wurden, so beobachtete Seebeck die Auslöschung des höheren Tons durch den tieferen:⁶⁹ „In diesem Falle heben sich nämlich die sämtlichen Impulse des tieferen gegen die abwechselnden Impulse des höheren auf, und es bleibt nur die andere Hälfte von denen des letzteren übrig, welche nun natürlich die tiefere Octave giebt.“ [sic]

Seebeck untersuchte auch Lochscheiben, deren Löcher nicht gleichmäßig angeordnet sind. Dabei stellte er fest, dass sich bei nur geringen Abweichungen von der Regelmäßigkeit, dennoch ein einzelner Ton einstellt. Variierte er den Abstand allerdings zu stark, so

⁶⁷Barkowsky (1996)

⁶⁸Seebeck (1841), Seite 419

⁶⁹Seebeck (1841), Seite 420

gelangte an das menschliche Ohr eine Folge von nicht isochronen Impulsen. Diese wurden in Folge derart interpretiert, als würden sie von mehreren Tönen stammen. Somit wären Kombinationstöne wahrnehmbar.⁷⁰ Von Kombinationstönen spricht man, wenn zusätzlich zu den im Klang vorhandenen Frequenzen f_1 und f_2 mit $f_1 < f_2$, die Summe $f_S = f_1 + f_2$ oder die Differenz $f_D = f_2 - f_1$ dieser Frequenzen gehört werden. Gelegentlich treten auch Kombinationstöne höherer Ordnung auf, z.B. der kubische Differenzton $f_3 = 2 \cdot f_1 - f_2$.

Barkowsky bezweifelte allerdings, dass Seebeck diese Töne tatsächlich gehört habe und belegte dies mit einem Hörbeispiel, das auf Seebecks Versuch beruhte.⁷¹

„Seebecks Aussage, daß Lochfolgen mit nur angenäherter Isochronie eine bestimmte Tonhöhe vermitteln, ist übertrieben. In Hörbeispiel 6 sind alle vier Klänge unterschiedlich. Der Anteil des 660 Hz Tons nimmt mit jedem Beispiel mehr ab. Georg Simon Ohm (1843; 1844) nimmt genau an dieser Seebeck'schen Sicht Anstoß, der zufolge ein Ton durch nichtisochrone Pulsfolgen entsteht.“

Ohm hatte bereits 1839 — ebenfalls in den Annalen der Physik und Chemie — Bemerkungen über Combinationstöne und Stöße gemacht. Er überprüfte dabei die Beobachtungen des finnischen Physikers Gustaf Gabriel Hällström (1775–1844), der festgestellt hatte, dass zwei Töne mit der Frequenz $r = nk$ und $s = mk$ mit $n, m, k \in \mathbb{N}$, sowie n, m teilerfremd, einen Ton mit der Frequenz des größten gemeinsamen Teilers $\text{ggT}(r, s) = k$ erklingen lassen.⁷² Hundert Jahre später konnten Wissenschaftler jedoch zeigen, dass dieser „unbedingte Combinationston“ genannte Ton kein Kombinationston im Sinne der Differenz- und Summationstöne ist (vgl. Abschnitt 4.4).

Dennoch war Ohms Entdeckung wichtig für die weitere (psycho)akustische Forschung. Denn das Phänomen dieser subjektiv hörbaren Töne, die nur durch Kombination von anderen Tönen entstehen, ist eine charakteristische Eigenschaft unseres Gehörs, die im Laufe der Zeit noch viele Wissenschaftler beschäftigen sollte.

⁷⁰Seebeck (1841), Seite 423

⁷¹Barkowsky (1996), Seite 216

⁷²Hällström (1832), Seite 440ff

Nach dem Erscheinen von Seebecks Artikel über seine akustischen Experimente 1841, veröffentlichte Ohm 1843 eine neue Theorie zur Tonhöhenwahrnehmung. Die mathematischen Erkenntnisse von Fourier (vgl. Formel (1)) hatten ihn dazu inspiriert.

Ohms Theorie besagte, dass jeder periodische Klang als Summe von einfachen harmonischen Schwingungen darstellbar sei, d.h. sich aus Sinustönen unterschiedlicher Amplitude zusammensetzen ließe. Das Ohr ordne diesem Klang in Folge eine Tonhöhe zu, die der Frequenz des tiefsten Tones entspräche.

Diese Ideen wurden in Wissenschafterkreisen nicht unbedingt mit Wohlwollen aufgenommen. Viele Kollegen vertraten andere Positionen. So schrieb etwa Röber, dass man Klänge durch diskrete Impulse („Binomien einer Verdichtung und Verdünnung“) wie bei der Sirene beschreiben könne.⁷³

Doch Ohm ließ sich davon nicht beirren und hielt an seiner Theorie fest. Diese warf allerdings die Frage auf, warum gerade ein bestimmter Ton, nämlich der tiefste, aus der Summe der sinusförmigen Teilschwingungen so stark hervortritt. Und hier hakte auch Seebeck ein, den Ohms Publikation zu heftigem Widerspruch veranlasste.

Er konnte der Idee, dass jeder Klang aus Sinustönen bestünde, nichts abgewinnen. Stattdessen zeigte er sich von der Tatsache überzeugt, dass auch andere Faktoren einen Einfluss auf die Tonhöhenwahrnehmung haben könnten:⁷⁴

„Ich sag: Unser Ohr empfindet den Eindruck einer periodischen Bewegung als Ton; alle Glieder, welche an dieser Periode Theil [sic] nehmen, können (wenigstens unter geeigneten Umständen) zur Stärke des Tones beitragen. Ohm sagt: Nur ein Glied bildet eigentlich den Hauptton, aber unser Ohr täuscht sich, indem es seine Beitöne mit zu ihm herüberzieht, und ihn selbst dadurch für stärker hält.“

Diese etwas aus der Luft gegriffen wirkende Erklärung von Ohm, wurde von Seebeck zu Recht kritisiert. Allerdings konnte auch Seebeck die Frage mit dem Grundton nicht beantworten. Die Lösung des Problems gelang jedoch einige Jahre später, als sich Helm-

⁷³zitiert nach Ohm (1843), Seite 514; Vermutlich sprach Ohm hier über den deutschen Gelehrten Friedrich Roeber.

⁷⁴Ohm (1843), S 361

holtz der Akustik zuwandte und die Zusammenhänge rund um Tonhöhe und Klangfarbe grundsätzlich geklärt wurden.

3.1.3 Helmholtz' Beitrag zur Tonhöhenwahrnehmung:

Der deutsche Physiker und Physiologe Hermann von Helmholtz (1821–1894) griff einige Jahrzehnte später Ohms Theorie wieder auf.

Bereits 1857 erläuterte er bei einem Vortrag die Idee, dass in der Cochlea eine Frequenzanalyse passiere.⁷⁵

„Das Ohr kann also, der Erfahrung nach, zusammengesetzte Luftbewegungen in Theile zerlegen. Unter zusammengesetzten Luftbewegungen haben wir bisher solche verstanden, die durch Zusammenwirkung mehrerer gleichzeitig tönender Körper entstanden waren. Da nun die Form der Tonwellen der verschiedenen musikalischen Instrumente verschieden ist, so wird es vorkommen können, dass die Schwingungsart der Luft, die ein solcher Ton im Gehörgange erregt, genau gleich ist der Schwingungsart, welche in einem anderen Fall von zwei oder mehreren anderen zusammenwirkenden Instrumenten im Gehörgange erzeugt wird. Wenn das Ohr im letzteren Falle die Bewegung in einzelne Theile zerlegt, wird es nicht umhin können, dasselbe auch im ersteren Falle zu thun, wobei der Ton nur aus einer Tonquelle her stammt. Und in der That geschieht dies.“ [sic]

Helmholtz 1863 erschienenen Buch *Die Lehre von den Tonempfindungen* wurde schon bald ein Standardwerk auf dem Gebiet der Akustik und speziell der Psychoakustik. Er unterschied darin strikt zwischen den Begriffen Klang und Geräusch (periodische, regelmäßige Luftschwingung vs. nicht periodische, unregelmäßige) und hielt explizit fest, dass sich ein Klang eindeutig durch die Attribute Stärke, Tonhöhe und Klangfarbe beschreiben ließe. Physikalisch gesehen also durch Amplitude, Frequenz und das Auftreten von Obertönen.⁷⁶

Mit Hilfe von Resonatoren (Kugel mit Öffnung, die genau eine Frequenz verstärkt und alle anderen dämpft), Stimmgabeln und Klaviersaiten untersuchte er die Obertöne und

⁷⁵von Helmholtz (1971) [1857], Seite 33-34

⁷⁶von Helmholtz (2003) [1863], Seite 19

die Zusammensetzung eines musikalischen Klanges genauer. Dabei erkannte er, dass bei einer am Klavier gezupften Saite nicht nur die Obertöne (Vielfache der Grundfrequenz), sondern auch die Subharmonischen (Teiler der Grundfrequenz) mitschwingen.

Er kam zu dem Schluss:⁷⁷

„Durch die angegebenen Erfahrungen wird nun auch der von G.S. Ohm aufgestellte und verteidigte Satz als richtig erwiesen, daß das menschliche Ohr nur eine pendelartige Schwingung der Luft als einfachen Ton empfindet, und jede andere periodische Luftbewegung zerlegt in eine Reihe von pendelartigen Schwingungen, und die diesen entsprechende Reihe von Tönen empfindet.“

Helmholtz beschäftigte sich auch mit Seebecks Kritik an Ohms Theorie und versuchte eine Erklärung zu finden, wieso der Grundton trotz gleicher bzw. geringerer Amplitude als die Obertöne so dominant zu hören ist. Er schrieb dies der Tatsache zu, dass das menschliche Gehör gewohnheitsmäßig die Obertöne nicht als Einzeltöne sondern als Klangfarbe wahrnehme.⁷⁸

Im Übrigen gelang es Helmholtz auch erstmals Summationstöne zu beobachten. Bisher waren als Kombinationstöne nur Differenztöne bekannt. Er schrieb deren Auftreten der Schnittstelle Trommelfell/Gehörknöchelchen zu.⁷⁹

3.1.4 Erkenntnisse der Neurologie im 20. Jahrhundert

Neue technische Innovationen führten zu Beginn des 20. Jahrhunderts zur Gründung neurologischer Institute. Die Fortschritte in der Elektrotechnik ermöglichten auch ein tieferes Verständnis über die Funktionsweise von Nervenzellen (vgl. Abschnitt 2.2.1).

So konnte beispielsweise gemessen werden, dass zwischen zwei Aktionspotentialen eines Neurons aus physiologischen Gründen stets eine Pause, die Refraktärzeit, eingehalten

⁷⁷von Helmholtz (2003) [1863], Seite 97

⁷⁸Turner (1977), Seite 20

⁷⁹von Helmholtz (2003) [1863], Seite 257ff

werden muss. Dadurch ist die Feuerfrequenz von Nervenzellen nach oben mit ca. 3000–4000 Hz begrenzt.⁸⁰

In den 30er Jahren beobachteten der amerikanische Psychologe Ernest Glen Wever (1902–1991) und sein Kollege Charles W. Bray bei Untersuchungen an einer Katze, dass sich die Frequenz eines Eingangssignals in der Feuerrate des Hörnervs widerspiegelte:⁸¹ „These facts show, then, that in the nerve as a whole the frequency of response is correlated with the frequency of stimulation. There was evidence also of a correspondence between the intensity of response in the nerve as a whole and the intensity of stimulation.“ Diese Arbeit lieferte die Grundlage für die Theorie des *Volleyprinzips*. Damit ist gemeint, dass Neuronen in Gruppen zusammenarbeiten, um auch höhere Frequenzen übertragen zu können.⁸²

„[...E]ven though a nerve fiber may not fire in response to each cycle of the sound wave, a period histogram may show that the nerve fiber transmits information about the sound wave by locking its activity to a particular phase of the input waveform. Since auditory nerve fibers may phase lock to highfrequency sound without responding to each cycle, it is possible that several fibers firing during different cycles of the stimulus and converging on a central target could provide the brain with input about each cycle of a highfrequency sound.“

Das Volleyprinzip besagt also, dass die Informationsübertragung nicht durch das Feuermuster einer einzelnen Nervenzelle hervorgerufen wird, sondern über das Feuermuster mehrerer Nervenfasern, die als Gruppe operieren. Vor dem Eintreffen des akustischen Signals kommt es zu ständigen spontanen Entladungen der Neuronen. Diese passieren rein zufällig bis zu 100 Mal pro Sekunde.⁸³ Das Auftreten von Entladungen in einem festen Zeitraum wird heutzutage etwa durch einen Poisson-Prozess⁸⁴ simuliert. Bei der

⁸⁰Roeckelein (1998), Seite 61

⁸¹Wever u. Bray (1930), Seite 345

⁸²Kandel u. a. (1991), Seite 491

⁸³Hartmann (1996), Seite 3496

⁸⁴Fricke (2011), Seite 103;

Ein Poisson-Prozess ist ein stochastischer Prozess, der einen poissonverteilten Zuwachs modelliert.

Übertragung einer Frequenz beginnen die Nervenzellen ihr ehemals spontanes Feuermuster aufeinander abzustimmen, um auch höhere Frequenzen als Gruppe übertragen zu können.

Dieses zeitlich versetzte Feuern erinnert an Militärsalven (engl. volley), daher der Name Volleyprinzip. Die biologische Voraussetzung für die Übertragung höherer Frequenzen ist hiermit also gegeben. Es zeigt sich allerdings bei höheren Frequenzen ab etwa 5 kHz ein Nachlassen der Synchronizität der Neuronen. Dies ist darauf zurückzuführen, dass es den Nervenzellen nicht möglich ist ein derart hochfrequentes Feuermuster mit zeitlich exakter Präzision auszuführen.⁸⁵

In den höchsten wahrnehmbaren Frequenzbereichen - im Umfang von etwa 2 Oktaven - ist es Menschen und anderen Säugetieren folglich kaum möglich einem Klang eine Tonhöhe zuzuordnen. So stellten Bizley und Walker bei Messungen an Katzen fest, dass die obere Grenze der Tonhöhenwahrnehmung etwa bei 5 kHz liegt.⁸⁶

3.1.5 Erkenntnisse der Psychoakustik im 20. und 21. Jahrhundert

In psychoakustischer Hinsicht gab es in der ersten Hälfte des 20. Jahrhunderts keine herausragenden neuen Erkenntnisse. Die Frage, welche Art von Wellen auf der Basalmembran erzeugt wird, war zu diesem Zeitpunkt immer noch ungelöst. Es gab unterschiedliche sich widersprechende Theorien, bis Békésy seine Forschungen zur Basalmembran publizierte. Es erkannte, dass er alle Theorien durch die Variation eines einzelnen Parameters beschreiben konnte:⁸⁷

„Although the handbooks especially were much occupied with pointing out the differences among the four theories, I reversed the question and tried to find their common features. The result was surprising: it turned out that it was possible to go continuously from one theory to the other merely by changing the numerical value of one parameter, namely, the volume elasticity.“

Die Abstände zwischen zwei Zuwächsen sind dabei exponentialverteilt. ("Erneuerungsprozess")

⁸⁵Hartmann (1996), Seite 3496

⁸⁶Bizley u. Walker (2010), Seite 4

⁸⁷von Békésy (1961), Seite 741

Békésy gelang es schließlich in seinen Experimenten den entscheidenden Parameter, die Elastizität der Basilarmembran, zu messen. Somit konnte er daraus schließen, welche der Theorien die richtige war. Es stellte sich heraus, dass es Wanderwellen sind, die sich auf der Basilarmembran fortpflanzen. Für seine Arbeit erhielt er 1961 den *Nobelpreis für Physiologie oder Medizin*.

Gegen Ende des 20. Jahrhunderts wurden bedeutende Fortschritte in der Erforschung des menschlichen Gehirns erzielt. Die technische Entwicklung moderner Untersuchungsmethoden (z.B. funktionelle Magnetresonanztomographie: ein bildgebendes Verfahren, das neuronale Aktivität im Gehirn auf Grund von Durchblutungsänderungen sichtbar macht) ermöglichte es, Hirnareale zu bestimmen, die bei gewissen Anregungen - wie etwa bei Änderung der Tonhöhe - aktiv sind.

Dabei stellte sich heraus, dass die Tonhöhenwahrnehmung nicht auf eine spezielle Gehirnregion beschränkt ist, sondern je nach Stimulus verschiedene Bereiche des auditorischen Cortex neuronale Aktivitäten aufweisen. Bizley und Walker formulierten diesen Sachverhalt folgendermaßen:⁸⁸

„Thus, in contrast to our expectation to find a universal pitch center within the auditory cortex where neurons respond invariantly to the perceived pitch of all sounds, there is no evidence to suggest that, instead, a network of pitch-sensitive regions in the auditory cortex exists to support a variety of periodicity judgments.“

Es dürften somit mehrere räumlich getrennt liegende Bereiche des Gehirns bei der Tonhöhenwahrnehmung zusammenwirken. Unterschiedliche Aspekte und Phänomene der Tonhöhenwahrnehmung erklären sich wahrscheinlich durch die Reizverarbeitung in verschiedenen Gehirnregionen und -strukturen. Viele Fragen zu dieser Thematik sind allerdings noch ungeklärt:⁸⁹

„Many questions remain concerning how auditory cortical neurons encode the perception of pitch. For instance, how are the harmonic template and autocorrelation representations of periodicity observed at the subcortical level transformed into a single spike code

⁸⁸Bizley u. Walker (2010), Seite 6

⁸⁹Bizley u. Walker (2010), Seite 7

for pitch by auditory cortical neurons? Is pitch, in fact, represented by single cortical neurons, or can level- and spectral-invariant representations of pitch only exist across populations of neurons? Finally, are neurons in different cortical regions specialized to extract different aspects of pitch, such as pitch chroma and height?”

Das 21. Jahrhundert wird dank neuer technischer Möglichkeiten (bildgebende Verfahren, hohe Rechenleistung bei Computern, Simulationsprogramme für dynamische Systeme, etc.) hoffentlich bald Antworten auf diese ungelösten Fragen liefern können.

3.2 Zusammenfassung und aktueller Forschungsstand zur Tonhöhenwahrnehmung

Im Wesentlichen haben sich durch die Forschungsergebnisse der Vergangenheit zwei mögliche Theorien zur Tonhöhenwahrnehmung herauskristallisiert. Das ist einerseits die Zeitbereich- und andererseits die Frequenzbereichsanalyse. In den folgenden Kapiteln sollen die beiden Theorien erläutert und näher untersucht werden. Das Residualtonhören als Phänomen der Tonhöhenwahrnehmung soll dabei als Kriterium dienen.

3.3 Das Residualtonhören als Phänomen der Tonhöhenwahrnehmung

Wenn ein Klang, der aus einem Grundton und den harmonischen Obertönen besteht, den Grundton nur sehr schwach enthält bzw. dieser sogar entfernt wird, so verarbeitet das Gehör den Klangeindruck in Bezug auf die Tonhöhe genauso, als wäre dieser Grundton vorhanden. Das bedeutet, dass die Tonhöhenempfindung des Klanges gleich bleibt. Man spricht dann vom Residualton und bezeichnet damit jene Tonhöhe, die dem Klang zugeordnet wird. Allein die Klangfarbe unterscheidet den vollständigen Klang inklusive Grundton, von demjenigen ohne Grundton.

In der Praxis wurde dieses Phänomen zunächst bei Instrumenten beobachtet (vgl. Abschnitt 3.1.2). So wurde beim Streichen der g-Saite einer Geige gemessen, dass die Ober-

töne viel stärker auftreten als der eigentliche Grundton *g* und auch beim Vokal *A* kam es zu denselben Ergebnissen.⁹⁰ Bei tiefen Instrumenten, fehlt der Grundton manchmal sogar ganz. So kann das Fagott auf Grund seiner zu geringen Größe bei tiefen Tönen den Grundton nicht stark genug abstrahlen. In diesem Fall bleibt der Grundton unter der Hörschwelle.⁹¹ In all diesen Situationen tritt ein Residualton derselben Frequenz an die Stelle des Grundtons.

Ein Spezialfall sind Instrumente wie beispielsweise Glocken, die ein nicht-harmonisches Spektrum aufweisen, wobei die wahrgenommene Tonhöhe in diesem gar nicht auftritt. Auch hier tritt eine Art Residualtonwahrnehmung („Schlagton“) anstelle eines Grundtones. Bei schlecht konstruierten („ugly“) Glocken, die kaum einen charakteristischen Klang aufweisen, gelingt dies auf Grund der Struktur des Obertonspektrums hingegen nicht.⁹² Es liegt die Vermutung nahe, dass der Glockenton vom Gehör aus dem Glockenspektrum extrahiert wird, indem ein annähernd harmonisches Teilspektrum daraus zur Residualtonbildung verwendet wird. Bei schlechten Glocken fehlt der Residualton folglich, da sich kein geeignetes harmonisches Teilspektrum finden lässt.

Später tauchte das Residualtonphänomen beim Telefonieren auf, als festgestellt wurde, dass die Übertragung des Frequenzbereiches unter 300 Hz nicht notwendig ist, um die Sprecherstimme richtig wahrzunehmen.⁹³

Das verwendete Frequenzband bei der Übertragung liegt ca. zwischen 500 und 4000 Hz.⁹⁴ Alle höheren oder tieferen Töne kommen beim Empfänger physikalisch gesehen nicht an. Trotzdem erklingen Stimmen, deren Grundton etwa bei 100 Hz (Männer) bzw. 200 Hz (Frauen)⁹⁵ liegt, beim Telefonieren in ihrer gewöhnlichen Tiefe und nicht höher. Die Einschränkung im Frequenzspektrum wirkt sich lediglich auf die Klangqualität aus.⁹⁶

Auch bei Lautsprechern, die tiefe Frequenzen beispielsweise im Klang eines Kontrabasses

⁹⁰Schouten (1940) Seite 287

⁹¹Fricke (2011), Seite 106

⁹²Schouten (1940), Seite 293

⁹³Schouten (1940), Seite 287

⁹⁴Nelken u. Ahissar (2006), Seite 6

⁹⁵Langner (2007), Seite 11

⁹⁶Terhardt (1982), Seite 25

nicht übertragen können, beeinflusst diese Tatsache nur die Klangfarbe, nicht aber die Tonhöhe.

Vorausschickend sei festgehalten, dass es der Wissenschaft bis heute noch nicht gelungen ist, im menschlichen Gehirn ein Zentrum für das Residualtonhören zu identifizieren — ebenso wenig wie ein Tonhöhenzentrum. Das bedeutet, dass bisher keine Neuronen gefunden werden konnten, die mit ihrem Feuermuster eine Residualtonfrequenz codieren.⁹⁷ Im folgenden Kapitel werden das Residualtonhören und zahlreiche Forschungsergebnisse aus diesem Bereich ausführlich dokumentiert.

⁹⁷Bizley u. Walker (2010), Seite 7

4 Das Residualtonphänomen

4.1 Der Residualton — eine Begriffsverwirrung

Auf Grund der unterschiedlichen Disziplinen und Forschungsansätze, die Wissenschaftler zum Thema Residualtonhören geführt haben, gibt es sehr viele verschiedene Begriffe dafür, wie zum Beispiel:

Unbedingter Combinationston	Ohm (1839)
Missing fundamental	Hoogland (1953)
Residuum	Meyer-Eppler u. a. (1959)
Residue	Schouten u. a. (1962); Ritsma (1962)
Periodenton(höhe)	Walliser (1969b)
Pitch of complex tones	Goldstein (1973)
Virtuelle Tonhöhe	Terhardt (1982)
Missing-F0 pitch	Tramo u. a. (2005)

Zwecks besserer Verständlichkeit wird in dieser Arbeit stets der Begriff Residualton verwendet, unabhängig davon, welchen Begriff die jeweiligen Autoren ursprünglich wählten.

4.2 Die Entdeckung des Residualtonphänomens

Erstmals ausführlich erforscht wurde das Residualtonphänomen von Jan Frederik Schouten (1938/40), dessen Artikel *The Residue revisited*⁹⁸ einen guten Überblick über die geschichtliche Entwicklung der Thematik gibt.

⁹⁸Schouten (1970)

1938 veröffentlichte er zunächst den Artikel *The perception of subjective tones*⁹⁹, worin er sich mit der Wahrnehmung von Tönen, die nicht im Frequenzspektrum auftreten, beschäftigte. Er nannte diese Töne subjektive Töne im Gegensatz zu den objektiven, die einer Frequenz im Klangspektrum entsprechen. Seine Forschungsarbeit handelte von Differenz-, Summations- und Residualtönen („residue“). Diese drei Phänomene eint, dass Töne gehört werden, die im Frequenzspektrum rein physikalisch nicht vorhanden sind. Während Summations- und Differenztonbeobachtungen schon früher von Ohm, Seebeck oder Helmholtz ausführlich untersucht worden waren, so etablierte sich das Residualtonhören als ein neues Forschungsgebiet.

Harvey Fletcher¹⁰⁰ führte 1924 Experimente mit Telefonleitungen durch, in denen Filter gewisse Frequenzen aus der Übertragung entfernten. Dabei stellte er fest, dass die wahrgenommene Tonhöhe des Vokals „ah“ gleich bleibt, selbst wenn der Grundton und die ersten sieben Obertöne herausgefiltert werden. Er vermutete die Ursache für diese Wahrnehmung in nichtlinearen Verzerrungen bei der akustischen Verarbeitung im Ohr und meinte, dass der Grundton als Differenzton der vorhandenen Obertöne entstünde:¹⁰¹ „Recent work on hearing has shown that the transmission mechanism between the air and the inner ear has a non-linear response which accounts for the so-called subjective tones. When the components of low frequency are eliminated from the externally impressed musical tone, they are again introduced as subjective tones before the sound reaches the nerve terminals.“

Gerrit Anthonie Hoogland beobachtete 1953 in einem Experiment, dass der Residualton auch dann hörbar ist, wenn die Obertonfrequenzen von weißem Rauschen¹⁰² überdeckt werden.¹⁰³ Im Gegensatz zu den im nächsten Abschnitt beschriebenen Messungen von Schouten stellte er sehr wohl eine Schwebung zwischen Residualton und einem Ton mit fast gleicher Frequenz fest. Daher war er wie Fletcher davon überzeugt, dass der

⁹⁹Schouten (1938)

¹⁰⁰Fletcher (1924)

¹⁰¹Fletcher (1924), Seite 427

¹⁰²Das weiße Rauschen enthält alle Frequenzen mit konstanter Leistungsdichte.

¹⁰³Hoogland (1953)

Residualton auf Grund von Verzerrungen im Ohr entstünde.

Später wurde diese These für den Residualton widerlegt. Dies deutet darauf hin, dass Hoogland nur einen Kombinationston und keinen Residualton wahrnahm (vgl. Abschnitt 4.4). Nichtlineare Verzerrungen können nicht für das Residualtonhören verantwortlich sein, wie u.a. Werner Meyer-Eppler et al. in ihren Untersuchungen feststellten. Sie präsentierten Probanden zunächst in einem Ohr ein 400 Hz-Signal und gleichzeitig im anderen Ohr ein 500 Hz-Signal („diplotisch“). Wenn man anschließend beide Frequenzen in beiden Ohren abspielt („amphotisch“), ist zusätzlich ein Residualton wahrnehmbar. Dieses Phänomen tritt auch bei sehr schwacher Lautstärke auf.¹⁰⁴

„Verringert man nämlich den Schalldruck der beiden Primärsignale des oben angeführten Beispiels bis nahe an die Hörschwelle, so bleibt dennoch der Umschlag der Empfindungsqualitäten beim Übergang von der diplotischen zur amphotischen Darbietungsweise voll erhalten. Es gibt keine Art von nichtlinearer Verzerrung, die mit diesem Sachverhalt verträglich wäre.“

4.3 Eigenschaften des Residualtons

4.3.1 Schoutens Forschungsergebnisse

1962 schrieben Jan Frederik Schouten et al. nach zahlreichen Experimenten zu diesem Thema eine ausführliche Charakterisierung des Residualtonphänomens im Artikel *Pitch of the Residue*:¹⁰⁵

„It was discovered when listening to a periodic pulse of repetition rate of 200 cps, the spectrum of which contained a score of Fourier components. In subjective sound analysis, only the lower harmonics can be heard separately with a loudness decreasing with increasing order. The higher harmonics are not individually perceptible, but are heard as a single sharp note with a pitch equal to that of the fundamental tone. [...]ts loudness depends on the number and intensity of the higher harmonics and is fairly independent

¹⁰⁴Meyer-Eppler u. a. (1959), Seite 72

¹⁰⁵Schouten u. a. (1962), Seite 1418

of those of the lower ones. [...] Beats of a residue and a pure tone of corresponding pitch do not occur. [...] Masking of a residue and a pure tone of equal pitch is not observed.“
Dadurch, dass beim Residualton keine Schwebungs- und Maskierungseffekte auftreten, ist Fletchers und Hooglands Theorie, dass der Residualton auf nichtlinearen Verzerrungen im Innenohr beruhe, widerlegt.

Bei Untersuchungen beobachteten Schouten et al. eine weitere interessante Eigenschaft. Ab einer Frequenz von ca. 3 kHz kann einem Residualton keine klare Tonhöhe mehr zugeordnet werden.¹⁰⁶ Manche Forscher sprechen daher von einem tonalen und einem atonalen Residualton.¹⁰⁷ Eine profunde Theorie für das Residualtonhören muss also diesen Umstand hinreichend erklären können.

Zwei weitere Aspekte der Arbeit von Schouten et al. sind von besonderer Bedeutung. Zum einen untersuchten sie, was sich bei einer Verschiebung des harmonischen Obertonspektrums um ein Frequenzintervall Δf ändert. In diesem Fall können die Obertonfrequenzen nicht mehr länger als Teil eines harmonischen Spektrums aufgefasst werden:¹⁰⁸
„As an empirical rule, it was found that the pitch of such a residue moves up and down in linear proportion to the center frequency of the constituent components. We refer to this as the first effect of pitch shift.“

In Folge ändert sich der Residualton ebenfalls, allerdings in weit geringerem Ausmaß als die harmonischen Obertöne.

Zum anderen beobachteten die Forscher die Veränderung des Residualtons bei Vergrößerung der Frequenzabstände der drei harmonischen Obertöne, wenn der mittlere Ton seine Frequenz behält. So kommt es zu einem Abfallen des Residualtons, wenn dieser Abstand zunimmt. Schouten et al. bezeichneten diese Beobachtung als *second effect of pitch shift*. Dieser Effekt tritt sogar noch stärker zutage als der erstgenannte.¹⁰⁹

¹⁰⁶Schouten u. a. (1962), Seite 1419

¹⁰⁷In Schouten u. a. (1962), Seite 1419 wurde etwa auf Egbert de Boer verwiesen.

¹⁰⁸Schouten u. a. (1962), Seite 1420

¹⁰⁹Schouten u. a. (1962), Seite 1420

Abschließend schlugen die Autoren als Erklärung für die auftretenden Effekte einen neuronalen Mechanismus in Form einer Verzögerungskette vor.¹¹⁰

Zu Schoutens Versuchen finden sich bei Barkowsky einige Hörbeispiele, aus denen unmittelbar der Schluss gezogen werden kann, „daß die Wahrnehmung von Residualtönen keine Folge von Auflösungsschwierigkeiten des Ohres für eng beeinanderliegende Frequenzen ist, und daß die Wahrnehmung von Residualtönen eine Folge der neuronalen Verarbeitung ist.“¹¹¹

4.3.2 Ritsmas Forschungsergebnisse

Roelof Johan Ritsma beschrieb in seinen beiden Artikeln *Existence Region of the Tonal Residue I* (1962) und *Existence Region of the Tonal Residue II* (1963) die genauen Entstehungsbedingungen für einen Residualton, welchen er mittels Modulation erzeugte. Ritsma erstellte aus seinen Daten eine Art detaillierte Landkarte für die Frequenzbereiche, in denen ein Residualton gehört werden kann.

Dadurch konnte er zeigen, warum es Hoogland in einigen seiner Experimente nicht gelungen war, einen Residualton zu hören. Hoogland befand sich mit seinem Versuchsaufbau außerhalb des Bereichs, in dem eine Residualtonwahrnehmung möglich ist.¹¹²

Auch Ritsma schloss sich der Meinung an, dass der Residualton keine Maskierungen oder Schwebungen mit anderen Tönen bildet und daher rein auf Grund der Basilarmembranerregungen entstehen müsse:¹¹³

„Thus a residue [...] must be interpreted as being generated at the location of the constituent frequencies on the basilar membrane. This necessitates the supposition that each area of the basilar membrane may give rise to sensations of widely different pitch, depending upon the pattern of stimulation.“

¹¹⁰Schouten u. a. (1962), Seite 1424

¹¹¹Barkowsky (1996), Seite 261

¹¹²Ritsma (1962), Seite 1227

¹¹³Ritsma (1962), Seite 1224

Kurz zusammengefasst beobachtete Ritsma, dass der Residualtonfrequenzbereich mit abnehmender Modulationstiefe¹¹⁴ schrumpft; sowie dass die absolute Frequenz der harmonischen Obertöne weniger entscheidend als deren relativer Abstand zueinander ist.¹¹⁵ Aus seinen Untersuchung zog er den Schluss, dass die Tonhöhenwahrnehmung, und somit auch das Residualtonhören, durch neuronale Prozesse, die erst nach der Basilarmembran stattfinden, ablaufen. Einen weiteren wichtigen Hinweis lieferte er mit der Beobachtung, dass über 5000 Hz kein Residualton mehr wahrgenommen wird. Ritsma machte das in diesem Frequenzbereich nicht mehr funktionierende Volleyprinzip dafür verantwortlich. Er propagierte infolge eine Art Autokorrelationsanalyse, indem er sich auf Licklider (vgl. Abschnitt 7.1) berief und schrieb:¹¹⁶

„[..]t may not be too farfetched to assume that the pitch extractor is made up of a great number of coincidence elements corresponding to specific delay times.“

Im zweiten oben genannten Artikel ergänzte Ritsma seine Arbeit noch um einige weiterführende Erkenntnisse in Bezug auf die nötige Anzahl der Obertöne und den Einfluss der Schall-Einhüllenden auf das Residualtonhören. So stellte er fest, dass ein Residualton nur dann gehört wird, wenn sich mindestens drei Obertonfrequenzen in dem im vorigen Artikel beschriebenen Residualtonfrequenzbereich befinden.¹¹⁷

„In taking a complex consisting of a greater number of components, the nerve fibers of a larger area of the basilar membrane will be activated. Consequently, more than one single integration bundle will be operative, but it might not be too far-fetched to assume that only one bundle needs to correspond to a frequency below the critical one, so as to be capable of bringing about a sensation of pitch. From this it would follow that, for a complex sound, a tonal residue can be established if, and only if, at least one integrating bundle is activated corresponding to a frequency below the critical one as found for a three-component complex.“

¹¹⁴ Die Modulationstiefe wird durch das Verhältnis von Nutzsignalamplitude zu Trägeramplitude erklärt.

¹¹⁵Ritsma (1962), Seite 1226–1227

¹¹⁶Ritsma (1962), Seite 1228

¹¹⁷Ritsma (1963), Seite 1245

4.3.3 Wallisers Forschungsergebnisse

Konrad Walliser bezeichnete den Residualton als Periodentonhöhe und publizierte 1969 seine Untersuchungsergebnisse zu diesem Thema.¹¹⁸ Er verwendete Klänge, die aus zumindest drei äquidistanten Sinustönen aufgebaut waren und unterschied dabei zwischen harmonischen und unharmonischen Spektren. Den Abstand zwischen den Frequenzen bezeichnete er als Modulationsfrequenz f_{mod} :

$$f_{mod} = f_3 - f_2 = f_2 - f_1$$

Im harmonischen Fall haben die Frequenzen f_1 , f_2 und f_3 einen gemeinsamen Teiler und lassen sich in der folgenden Gestalt formalisieren:

$$f_i = n f_{mod} \quad \text{mit } n \in \mathbb{N} \text{ und } i \in \{1, 2, 3\}$$

Im unharmonischen Fall ist dies nicht möglich.

Bei seinen Untersuchungen verschob Walliser einen Tonkomplex mit der mittleren Frequenz $f_m = 2$ kHz und Modulationsfrequenz $f_{mod} = 300$ Hz entlang der Frequenzachse. Das bedeutet, dass er die harmonische Ausgangslage der Sinustöne in einen unharmonischen Zustand überführte. Er beobachtete:¹¹⁹

„Wie ersichtlich, ist die Änderung Δf_v streng proportional zu der Verschiebung Δf : $\Delta f_v = \text{const } \Delta f$. Die Periodentonhöhe [Anm.: der Residualton] nimmt in Abhängigkeit von Δf sägezahnförmig zu und ab. für Frequenzverschiebungen $\Delta f = \pm f_{mod}$ erhält man wie für den Ausgangsfall $\Delta f = 0$ einen harmonischen Fall und nimmt die gleiche Tonhöhenempfindung wahr. Zwischen den Werten $\Delta f = 0$ und $|\Delta f| = f_{mod}$ tritt ein Sprung der Periodentonhöhe auf, im untersuchten Beispiel um etwa 20%.“

Diese Resultate knüpften unmittelbar an Schouten et al. an (vgl. Abschnitt 4.3.1, *first effect of pitch shift*). In weiterer Folge beschäftigte Walliser sich auch mit dem *second*

¹¹⁸Walliser (1969b)

¹¹⁹Walliser (1969b), Seite 322; Anm.: Zwecks einheitlicher Darstellung wird in dem Zitat Δf anstatt des von Walliser verwendeten δf verwendet.

effect of pitch:¹²⁰ „Die Periodentonhöhe steigt mit wachsendem f_{mod} in treppenartigen Sprüngen (etwa 20% Änderung bei einem Sprung) mit ebenen Treppenstufen an; die Kurve schlängelt sich dabei um die Periodentonhöhe für den harmonischen Fall [...]“

Insgesamt beobachtete Walliser im unharmonischen Fall eine Abweichung des Residualtons um bis zu 10% im Vergleich zur Modulationsfrequenz:¹²¹

„Die Periodentonhöhe f_v kann also bei beliebigen äquidistanten Linienspektren bis zu etwa $\pm 10\%$ (mehr als ein Halbtonschritt) von der Modulationsfrequenz $f_{mod} = \frac{1}{T}$ abweichen.“

Walliser fasste seine Ergebnisse wie folgt zusammen:¹²²

„Die wichtigste Schallgröße ist die Hüllkurvenperiode $T = \frac{1}{f_{mod}}$ des akustischen Signals. Eine Verschiebung des Tongemisches aus dem harmonischen in den unharmonischen Fall kann aber noch Änderungen der Periodentonhöhe bis zu $\pm 10\%$ bewirken.“

Das bedeutet, dass Walliser ebenso wie Ritsma (vgl. Abschnitt 4.3.2) bei der Residualtonwahrnehmung die Abstände der Obertöne für relevanter als deren absolute Frequenz hält.

In einem weiteren Artikel *Zur Unterschiedsschwelle der Periodentonhöhe*¹²³ berichtete Walliser, dass der kleinst mögliche wahrnehmbare Frequenzunterschied („Unterschiedsschwelle“) bei Residualtönen nicht derselbe sei, wie beim Sinuston gleicher Frequenz. Die Schwelle liege, vor allem bei tiefen Tönen, klar darunter. Das ließe sich damit erklären, dass beim Residualton die Unterschiedsschwellen der einzelnen Obertöne eine entscheidende Rolle spielten. So sei beim Residualton ein Unterschied bereits dann wahrzunehmen, wenn auch nur einer der Obertöne über seiner Unterschiedsschwelle liege.

Die Unterschiedsschwelle des Residualtons ist somit stark an die Unterschiedsschwellen der Obertöne gekoppelt und nicht an die Unterschiedsschwelle des Sinustones gleicher

¹²⁰Walliser (1969b), Seite 324

¹²¹Walliser (1969b), Seite 324

¹²²Walliser (1969b), Seite 328

¹²³Walliser (1969a)

Frequenz.¹²⁴

4.3.4 Weitere Ergebnisse

Meyer-Eppler et al.¹²⁵ untersuchten, wie sich Melodien verändern, wenn zunächst der Grundton und nach und nach auch immer mehr Obertöne entfernt werden. Dabei stellten die Wissenschaftler fest, dass die Erkennung der Melodie durch das Weglassen kaum erschwert wird, allerdings die Bestimmung der Oktavlage ab einem bestimmten Punkt nicht mehr möglich ist. Sie interpretierten den Residualton als Quotiententon der Obertöne.¹²⁶ Diese Bezeichnung ist ein wenig irreführend, da sie an die Kombinationstöne Summations- und Differenzton erinnert. Der Residualton ist jedoch kein Kombinations- ton (vgl. Abschnitt 4.4).

Houtsma und Goldstein¹²⁷ spielten Probanden zwei nebeneinander liegende Frequenzkomponenten vor und untersuchten ob diese einen Residualton wahrnehmen. Zuerst wurde das Signal diotisch präsentiert, also beide Signalanteile in beiden Ohren; anschließend dichotisch, jeweils eine Frequenzkomponente des Signals in einem Ohr. Es stellte sich heraus, dass in beiden Fällen ein Residualton zu hören war. Daraus lässt sich schließen, dass das Residualtonhören nicht durch Mechanismen oder Abläufe in der Cochlea ausgelöst wird. Denn dort steht dem jeweiligen Ohr nur eine Frequenz zur Verfügung, aus der sich noch kein Residualton ableiten lässt. Somit dürfte kein Residualton hörbar sein. Dass er dennoch gehört wird, deutet darauf hin, dass er erst auf einer höheren Ebene der Hörbahn gebildet wird.

Abschließend ist noch festzuhalten, dass das Residualtonhören kein rein menschliches Phänomen ist. Auch bei anderen Säugetieren wurde diese Fähigkeit bereits festgestellt. So konnten Heffner und Whitfield¹²⁸ 1976 bei Versuchen mit zwei geschulten Katzen

¹²⁴Walliser (1969a), Seite 335

¹²⁵Meyer-Eppler u. a. (1959), Seite 71

¹²⁶Meyer-Eppler u. a. (1959), Seite 74

¹²⁷Houtsma u. Goldstein (1972)

¹²⁸Heffner u. Whitfield (1976)

zeigen, dass diese ebenfalls eine Art Residualton hören. Den Tieren gelang die Residualtonwahrnehmung bis zu einer Frequenz von ca. 6000 Hz, wobei es mit wachsender Frequenz für die Tiere zunehmend schwieriger wurde.

Whitfield¹²⁹ führte einige Jahre später dieses Experiment erneut durch. Dieses Mal allerdings mit Katzen, denen Teile des Hörcortex entfernt worden waren. Nur wenige der Tiere (2 von 13) waren danach noch fähig, fallende und steigende Tonfolgen voneinander zu unterscheiden. Weiters stellte sich heraus, dass die Katzen zwar noch unterschiedliche Frequenzen erkennen konnten, die Eigenschaft des Residualtonhörens aber verloren hatten.¹³⁰

Houtgast¹³¹ konnte zeigen, dass sogar ein einziger Oberton eine Residualtonwahrnehmung hervorrufen kann, wenn das Signal-Rausch-Verhältnis entsprechend gewählt wird. Für das Experiment verwendete er rosa Rauschen¹³² als Hintergrundgeräusch. Er beobachtete, dass es den 50 Probanden leichter fiel aus harmonischen Obertönen einen Residualton heraus zu hören und diesen mit einem gegebenen 200 Hz-Ton (bzw. dessen Obertonspektrum) zu vergleichen (Frequenz $\pm 3\%$), wenn das Hintergrundrauschen präsent war. Das Rauschen unterstützte also die Residualtonwahrnehmung.

Erstaunlicherweise konnten die Versuchsteilnehmer sogar aus einem einzelnen Oberton eine Art Residualtonempfindung ableiten und diese mit dem fix gegebenen 200 Hz-Ton vergleichen. Dies ist insofern bemerkenswert, da ein einzelner Sinuston im Obertonspektrum zahlreicher Grundtöne auftritt und es keinen speziellen Grund gibt, einen dieser Töne zu bevorzugen. Houtgast schreibt diese Wahrnehmung der Aufmerksamkeit der Probanden zu, die auf einen speziellen Frequenzbereich gelenkt wird.¹³³

„It is conceivable that a subject's attention during the experiment is being focussed on

¹²⁹Whitfield (1980), Seite 646

¹³⁰Auch wenn das Ergebnis von wissenschaftlicher Relevanz sein mag, möchte ich mich von derartigen Tierversuchen distanzieren.

¹³¹Houtgast (1976)

¹³²Das rosa Rauschen enthält alle Frequenzen mit logarithmisch sinkende Leistungsverteilung bei zunehmender Frequenz.

¹³³Houtgast (1976), Seite 407

that particular pitch region, and that of the potential subharmonic pitches of a pure-tone signal, the one around 200 Hz is strongly favored.“

Aus Houtgasts Experimenten ist die Schlussfolgerung zu ziehen, dass ein Sinuston neben der naheliegenden stets auch noch andere mögliche Tonhöhenempfindungen auslösen kann. Die Tonhöhe eines Klangs wird somit in gewissen Situationen je nach Aufmerksamkeit des Zuhörers, Hintergrundgeräuschen oder anderen Einflüssen unterschiedlich interpretiert. Houtgast sprach von „potential pitches“,¹³⁴ wobei je nach Situation eine davon realisiert werde. Diese Idee der Tonhöhe als subjektive Wahrnehmungsgröße wurde bereits im Abschnitt über die Tonhöhendefinition diskutiert (vgl. Abschnitt 1.2). Da Houtgasts Untersuchungen teilweise auf einer geringen Anzahl von Probanden beruhen ($n=3$),¹³⁵ wäre es wünschenswert seinen Versuchsaufbau erneut mit einer größeren Gruppe zu realisieren.

4.4 Residualton versus Kombinationston

In diesem Abschnitt sollen noch einmal kurz die wesentlichen Unterschiede zwischen Residualtönen und Kombinationstönen herausgearbeitet werden. Der Begriff Kombinationston umfasst im Folgenden sowohl Summationstöne, als auch Differenztöne.

Wie in Abschnitt 3.1.2 erwähnt, erkannte Ohm bereits, dass es sich hierbei um zwei unterschiedliche Phänomene handelt. Daher bezeichnete er Kombinationstöne als „bedingte Combinationstöne“ und Residualtöne als „unbedingte Combinationstöne“.¹³⁶

Zahlreiche Versuche mit Probanden zeigten, dass ein Residualton im Gegensatz zu einem Kombinationston keine große Lautstärke braucht.¹³⁷ Somit ist es naheliegend für diese beiden Klangphänomene zwei verschiedene Entstehungsmechanismen anzunehmen. Während Kombinationstöne auf Nichtlinearitäten in der Cochlea zurückgeführt werden,¹³⁸ gibt es beim Residualton Hinweise auf neuronale Autokorrelation (vgl. Abschnitt

¹³⁴Houtgast (1976), Seite 408

¹³⁵Houtgast (1976), Seite 406

¹³⁶Ohm (1839), Seite 463ff

¹³⁷Barkowsky (1996), Seite 234

¹³⁸Barkowsky (1996), Seite 234

6.2).

Kombinationstöne haben außerdem die Eigenschaft, dass sie im Ohr selbst wahrgenommen werden und nicht den Anschein erwecken, der Schallquelle im umgebenden Raum zu entspringen. So kann es passieren, dass ein Kombinationston in beiden Ohren unterschiedlich laut gehört wird.¹³⁹

Zur Illustration werden abschließend die folgenden beiden Residualton-Situationen verglichen:¹⁴⁰

1. 200 Hz, 600 Hz, 1000 Hz, etc.
2. 400 Hz, 800 Hz, 1200 Hz, etc.

In beiden Fällen ist der Abstand zwischen den Frequenzen jeweils 400 Hz. Wäre der Residualton ein Differenzton, so müsste in beiden Fällen ein 400 Hz-Ton gehört werden. Während im zweiten Fall tatsächlich 400 Hz wahrgenommen werden, so ruft das erste Beispiel eine Wahrnehmung von 200 Hz hervor. Der Residualton kann somit kein Differenzton sein.

4.5 Schlussfolgerungen für die Tonhöhenwahrnehmung

Aus der Beobachtung des Residualtonphänomens folgt insbesondere, dass das Ohmsche Gesetz modifiziert werden muss. Das Ohr zerlegt einen Klang zwar in seine Frequenzanteile. Allerdings können dabei sowohl einzelne Fourierkomponenten als Tonhöhe wahrgenommen, oder auch mehrere zu einer einzigen Tonhöhenwahrnehmung zusammengefasst werden, wie z.B. im Fall des Residualtons.

Diese Beobachtung soll nun dazu dienen, verschiedene Tonhöhenwahrnehmungsmodelle auf ihre Plausibilität hin zu überprüfen.

¹³⁹Barkowsky (1996), Seite 234

¹⁴⁰Schouten (1940), Seite 290

5 Analyse im Frequenzbereich vs. Analyse im Zeitbereich

Historisch gesehen reduzierten sich die zahlreichen Theorien zur Tonhöhenwahrnehmung im Laufe der Zeit nach und nach im Wesentlichen auf zwei Modelle. Einerseits die Analyse im Frequenzbereich, auch Ortstheorie genannt, und andererseits die Analyse im Zeitbereich, auch Periodizitätstheorie genannt. In diesem Kapitel werden die Grundlagen der beiden Theorien vorgestellt und auf ihre Tauglichkeit zur Erklärung des Residualtonhörens untersucht.

5.1 Ortstheorie/Analyse im Frequenzbereich

5.1.1 Ablauf der Analyse im Frequenzbereich

Die Ortstheorie basiert auf der Erkenntnis, dass auf der Basilarmembran eine Frequenzanalyse stattfindet und sich diese Tonotopie später auf allen Ebenen der Gehörbahn bis zum Hörcortex widerspiegelt. Alle weiteren Analysen im Gehirn würden in Folge auf der Frequenzanalyse in der Cochlea basieren.

Ein Vertreter der Ortstheorie, der sich mit Residualtönen beschäftigte, war Gerrit Anthonie Hoogland.¹⁴¹ Mit seinen Versuchen wollte er herausfinden, welche der beiden Theorien, die Zeit- oder die Ortstheorie, sich experimentell untermauern ließe. Dabei stellte er fest, dass auch Patienten mit high tone loss, die die Obertöne gar nicht wahrnehmen konnten, einen Residualton hörten. Auch bei Maskierung der Obertöne trat der Residualton auf. Daraus zog Hoogland den Schluss, dass die Residualtonwahrnehmung nicht auf der Wahrnehmung der Obertöne beruhe und durch nichtlineare Verzerrung entstünde. Diese Schlussfolgerung teilten andere nicht. Hooglands Thesen wurden größtenteils widerlegt (vgl. Abschnitt 4.1).

¹⁴¹Hoogland (1953)

5.1.2 Phänomene, die nicht durch Frequenzbereichsanalyse erklärt werden können

Zunächst stellt sich die Frage, ob die biologischen Voraussetzungen für eine Frequenzanalyse auf der Basilarmembran in derart feiner Auflösung gegeben sind, sodass die Tonhöhenwahrnehmung rein durch die Ortstheorie erklärt werden kann. Im Zuge von Békésys Forschungen stellte sich heraus, dass die Auflösung bei weitem nicht so fein wie erwartet war.¹⁴²

Der nächste Widerspruch besteht in der Tatsache, dass einem einzelnen Sinuston wesentlich schwerer eine Tonhöhe zugeordnet werden kann als einem Ton mit Grund- und Obertönen. Dies dürfte laut Ortstheorie nicht passieren, da im ersten Fall genau die eine der Frequenz zugehörigen Stelle in der Basilarmembran angesprochen wird. Horst-Peter Hesse schrieb dazu:¹⁴³

„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.“

Schließlich ist das Residualtonphänomen nicht allein durch die Ortstheorie erklärbar. Denn ein Klang, welcher die Grundtonfrequenz physikalisch nicht enthält, könnte keine Erregung an der entsprechenden Stelle der Basilarmembran auslösen. Der Residualton dürfte somit nicht auftreten.¹⁴⁴

Terhardt versuchte diese Probleme zu lösen, indem er die Theorie um einen Mustererkennungsprozess erweiterte.

¹⁴²Schouten (1970), Seite 45

¹⁴³Hesse (2003a), Seite 134–135

¹⁴⁴Schouten u. a. (1962), Seite 1420

5.1.3 Modifikation der Ortstheorie durch Terhardt

Ernst Terhardt schlug eine Modifikation der Ortstheorie in Form eines nachgeschalteten Mustererkennungsprozesses vor, um Residualtöne zu erklären. Zunächst werde im Ohr eine Frequenzanalyse durchgeführt und anschließend auf Grund des Frequenzabstandes der Obertöne eine Gestalt abgeleitet. Dieses Gestaltattribut sei dann der Residualton, bzw. die „Virtuelle Tonhöhe“, wie Terhardt dieses Phänomen allgemein bezeichnete. Mit seinen Experimenten bestimmte er, welche Obertöne dafür ausschlaggebend sind:¹⁴⁵

„Die Virtuelle Tonhöhe [Anm.: der Residualton] kann von geübten Versuchspersonen unter geeigneten Versuchsbedingungen bereits bei Darbietung von nur zwei Harmonischen wahrgenommen werden. [...] Am besten geeignet sind die ersten sechs Harmonischen.“

Die Fähigkeit diese Obertöne zu erkennen, führte Terhardt auf erlerntes Verhalten im Umgang mit Sprachlauten zurück:¹⁴⁶

„Die Sprache enthält natürlicherweise harmonische Klänge, nämlich die stimmhaften Sprachlaute, und wie oben bereits erwähnt, zerlegt das periphere Gehör solche Klänge innerhalb gewisser Grenzen in ihre Teiltöne. Der zentrale Verarbeitungsmechanismus des Gehörs wird andauernd vor die Aufgabe gestellt, die entsprechenden Spektraltonhöhen und damit auch die zwischen ihnen existierenden Intervalle auf ihren Informationswert hinsichtlich der sprachlichen Kommunikation zu untersuchen. Dabei kann es nicht ausbleiben, daß die betreffenden Intervalle vom Hörsystem im Laufe der Zeit als feste Parameter abgespeichert und danach beim aktiven Verarbeitungsprozess von Klangsignalen wieder verwendet werden.“

In weiterer Folge präsentierte er einen Algorithmus um Residualtöne mit Hilfe der Subharmonischen der Obertöne zu ermitteln:¹⁴⁷

Zunächst wird eine Spektralanalyse durchgeführt, bei der der Klang in die Frequenzkomponenten aufgespalten wird. Anschließend werden die Frequenzen all jener Töne bestimmt, die nicht durch Maskierung unwichtig werden. Da der Bereich von 500 Hz bis

¹⁴⁵Terhardt (1976), Seite 131

¹⁴⁶Terhardt (1982), Seite 33

¹⁴⁷Terhardt (1982), Seite 27

1500 Hz besonders stark in die Residualtonbildung eingeht, werden Obertonfrequenzen, die in diesen Frequenzbereich fallen, speziell gewichtet.

Zum Abschluss gilt es, für jede auf diese Weise ermittelte Frequenz all jene Obertonreihen zu finden, bei denen diese Frequenz eine der Positionen 2 bis 10 einnimmt. Diesen Vorgang nannte Terhardt „Subharmonische Koinzidenzdetektion“.

Hat man dieses Verfahren für die Frequenzkomponenten durchgeführt, so vergleicht man die für die jeweilige Frequenz gefundenen Obertonreihen miteinander. Diejenige Obertonreihe, welche am häufigsten in dieser Auflistung vorkommt, liefert den Residualton.

Nachteilig ist bei diesem Konzept, dass es nur für exakte Frequenzen funktioniert. Bei leichter Verstimmung kann diese Methode das Residualtonhören nicht mehr erklären. Terhardt war dieser Umstand bewusst. Er meinte, dass kleine Abweichungen vernachlässigbar wären:¹⁴⁸

„Es wird angenommen, daß die in den realisierten Akkordtönen unvermeidlich auftretenden Grundfrequenzabweichungen von der Normalstimmung so klein sind, daß sie für die Bildung der Virtuellen Tonhöhe bzw. der Grundtöne keine wesentliche Rolle spielen.“

Andere Forscher haben später versucht, diese Einschränkung durch die Einführung einer Frequenztoleranz bzw. Unschärfe zu beheben (vgl. Abschnitt 7.3.2).

Zusammengefasst basiert Terhardts Tonhöhenwahrnehmungstheorie somit auf einer Frequenzanalyse in Kombination mit einem erlernten Gestaltattribut, welches die Residualtonbildung ermöglicht.

¹⁴⁸Terhardt (1982), Seite 39

5.2 Periodizitätstheorie/Analyse im Zeitbereich

5.2.1 Ablauf der Analyse im Zeitbereich

Der Schweizer Mathematiker Leonhard Euler (1707–1783) beschäftigte sich im Zuge seiner Untersuchungen zur musikalischen Konsonanz und Dissonanz auch intensiv mit der Tonhöhe. Er stellte die Hypothese auf, dass die Tonhöhe durch ein rhythmisches Anstoßen des Trommelfells im Ohr erzeugt werde, und entwickelte daraus ein Konsonanz-Dissonanz-Modell für die Intervalle. Mittels Punktreihen veranschaulichte Euler die Impulse des Schalles.¹⁴⁹ Das zeitliche Zusammentreffen mehrerer Impulse unterschiedlicher Frequenzen gibt in Folge Aufschluss über die Wahrnehmung von Konsonanz bzw. Dissonanz.

Eulers Theorie bestand somit darin, dass Töne rhythmische Schwingungen im Ohr auslösen. Je mehr Pulse bei zwei Tönen zusammenfielen, desto konsonanter klangen sie.

1899 griff der deutsche Psychologe Theodor Lipps (1851–1914) diese Idee auf und entwickelte seine Mikrorhythmen-Theorie. Er meinte, dass die Nervenimpulse in der akustischen Verarbeitung den Rhythmus der Klangschwingungen codierten.¹⁵⁰

Die moderne Formulierung der Zeitbereichsanalyse spricht noch etwas präziser von der Codierung der Periodendauer des Schalls im Feuermuster der Neuronen des Hörnervs. Gegen diese Theorie sprach ursprünglich, dass die Feuerrate eines Neurons nach oben begrenzt ist. Höhere Frequenzen könnten auf diese Weise also nicht erreicht werden. Mit der Entdeckung des Volleyprinzips durch Glen Wever (vgl. Abschnitt 3.1.4) wurde dieses Problem behoben. Nicht einzelne Neuronen codieren die Periodendauer, sondern eine Gruppe von Neuronen, indem diese synchron versetzt feuern. Dadurch ist neurologisch erwiesen, dass auch höhere Töne in Form ihrer Periodendauer übermittelt werden können. Der zeitliche Abstand im Feuermuster einer Nervenfasers beinhaltet diese Information.

¹⁴⁹Übersetzung von Eulers *Tentamen Novae Theoriae Musicae* in Smith (1960), Seite 77

¹⁵⁰Hesse (2003b), Seite 122

Es bleibt die Frage, ob das menschliche Gehör überhaupt fähig ist, minimale Unterschiede im Zeitbereich auszuwerten. Mehrere Hinweise sprechen dafür, dass dies möglich ist. Ein Beispiel wäre das Richtungshören. Beim Richtungshören wird durch Laufzeitunterschiede des Schalls zwischen den beiden Ohren die Position der Schallquelle ermittelt. Dabei reichen Unterschiede von einigen Mikrosekunden (Langner¹⁵¹) bzw. $30 \cdot 10^{-9}$ sec (Hesse¹⁵²).

Hesse nannte als weiteres Beispiel die binauralen Schwebungen.¹⁵³ Bei einer Schwebung werden beiden Ohren gleichzeitig zwei Frequenzen präsentiert, die sich nur minimal voneinander unterscheiden. Es kommt beim Versuchsteilnehmer zu dem Sinneseindruck, als wäre nur eine Frequenz präsent, die mit einer bestimmten Periode lauter und leiser wird. Bei einer binauralen Schwebung wird jedem Ohr nur eine der beiden Frequenzen präsentiert. Dann entsteht beim Versuchsteilnehmer der Eindruck, als würde die Schwebung rund um den Kopf wandern. Man nennt dieses Phänomen auch Drehton:¹⁵⁴

„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.“

Die Voraussetzungen für feine Unterschiedsmessungen im Zeitbereich sind daher offenbar biologisch vorhanden. Die Beobachtungen lassen darauf schließen, dass die zeitliche Analyse von Tönen tatsächlich stattfindet, wie Hesse schrieb:¹⁵⁵

„Seit Mitte des 20. Jahrhunderts hat die neurophysiologische Forschung nach und nach immer mehr Beweise dafür gefunden, dass die Zeitstruktur des Schalles in Gestalt reiz-synchroner Nervenimpulse im unteren Teil des Hörnervensystems tatsächlich erhalten bleibt [...].“

¹⁵¹Langner (2007), Seite 10

¹⁵²Hesse (2003b), Seite 128

¹⁵³Hesse (2003b), Seite 129

¹⁵⁴Hesse (2003b), Seite 129

¹⁵⁵Hesse (2003b), Seite 128

5.2.2 Phänomene, die nicht durch Zeitbereichanalyse erklärt werden können

Die Cochleae der beiden Ohren sind naturgemäß nicht exakt gleich. Dadurch tritt bei der Tonhöhenwahrnehmung von Sinustönen der Diplacusis-Effekt auf:¹⁵⁶

„Diplacusis can be observed by using headphones that send a standard tone to one ear and then a matching tone to the other. When the subject matches the pitch of the standard by tuning the frequency of the matching tone, the tones in the two ears are found to have different frequencies.“

Weiters gibt es Beobachtungen, dass Sinustöne bei zunehmender Intensität¹⁵⁷ als höher (bei hoher Frequenz) oder tiefer (bei tiefer Frequenz) wahrgenommen werden. Dies bedeutet, dass die Lautstärke auch einen Einfluss auf die wahrgenommene Tonhöhe hat. Für diesen Effekt, manchmal auch *Stevens rule* genannt, liefert die Zeittheorie, die von der Lautstärke unabhängig operiert, keine Erklärung:¹⁵⁸

„Because the basilar membrane is a nonlinear neuromechanical system, the excitation pattern moves in the direction that agrees with Stevens rule for high frequencies, which actually tends to be the dominant effect perceptually. However, there is a problem because the peak of the excitation pattern shifts too much. It may shift by a place difference corresponding to almost an octave, whereas the pitch normally shifts by only a few percent.“

Hartmann erwähnte in seinem Artikel auch einen möglichen Lösungsvorschlag, die neuronale Autokorrelation, die im nächsten Abschnitt ausführlich vorgestellt wird:¹⁵⁹

„[.W]e know that pitch cannot be determined by the period of the neural spike train. This period is too rigidly tied to the period of stimulus to accomodate the pitch-shift effects. However, pitch could be determined by a derived quantity such as the neural autocorrelation function[...].“

¹⁵⁶Hartmann (1996), Seite 3494

¹⁵⁷ Anm.: Hartmann schrieb "intensity", also Schallintensität, bezog sich aber auf eine Graphik mit der Einheit db SPL

¹⁵⁸Hartmann (1996), Seite 3495

¹⁵⁹Hartmann (1996), Seite 3496

5.3 Verknüpfung der beiden Theorien (Frequenz- und Zeitbereichsanalyse)

Während früher noch Wissenschaftler ausschließlich eine der beiden Theorien befürworteten, so hat sich in den letzten Jahrzehnten ein weitgehender Konsens gebildet, dass nur durch das Zusammenwirken der beiden Analysen alle Phänomene des menschlichen Gehörs erklärt werden können. Durch die neuen Erkenntnisse der Hirnforschung und Neurologie (vgl. Langner in Abschnitt 7.1.2) gelang es insbesondere die Arbeitsweise der zeitlichen Analyse-Mechanismen besser zu verstehen und ihnen eine zentrale Rolle in der Tonhöhenwahrnehmung und Erklärung des Residualtonphänomens zuzuschreiben. In den weiteren Kapiteln wird die neuronale Verarbeitung einer speziellen Betrachtung unterzogen und ihr Potential zur Erklärung der Tonhöhenwahrnehmung analysiert.

Die Entdeckung des Volleypinzips (vgl. Kapitel 3.1) und Erkenntnisse über die Funktionsweise von Nervenzellen zeigten, dass biologisch gesehen eine Analyse im Zeitbereich möglich ist. Hesse formulierte in seinem Buch *Musik und Emotion* einen möglichen Ablauf für eine kombinierte Orts- und Frequenzanalyse:¹⁶⁰

„Ein komplex-periodischer Schall erzeugt im Innenohr eine örtlich aufgefächerte Erregungsverteilung. Der an den verschiedenen Orten des Cortischen Organs generierte 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 bedeutet, dass eine Zeitbereichsanalyse bis etwa 5 kHz funktionieren würde. In höheren Frequenzbereichen wäre nur eine Ortsanalyse möglich. Daher funktioniert in diesen Regionen die Tonhöhenwahrnehmung stark eingeschränkt.

Im weiteren Ablauf befürwortete Hesse Terhardts Theorie, dass „die Tonhöhenwahrneh-

¹⁶⁰Hesse (2003b), Seite 130

mung [...] auf einem Mustererkennungsprozess [beruht], der die örtliche und zeitliche Gliederung des Nervenimpulsmusters auswertet.“¹⁶¹

Durch das Zusammenfallen mehrerer Impulse, die verschiedenen Orten zugehören, ließe sich auf die Periode der Schallschwingung schließen. Diese zeitlichen Strukturen könnten in unteren Zentren der Hörbahn, vermutlich in den Cochleariskernen im Hirnstamm, erfasst werden.

Hesse zufolge wirft die Kombination von Ort- und Zeitbereichanalyse ein neues Licht auf die Unterschiede zwischen den Attributen Chroma und Tonhelligkeit.¹⁶² Das Chroma tritt grundsätzlich nur bei periodischen Schwingungen in Erscheinung und bei hohen Tönen fällt die Zuordnung schwer. Diese beiden Eigenschaften deuten auf einen engen Zusammenhang zwischen Chroma und Zeitbereichanalyse hin. Ohne eine periodische Klangstruktur ist die Wahrnehmung des Chromaattributes nicht möglich, daher ist die periodische Schwingung eine Grundvoraussetzung. Die obere Grenze von etwa 4000 Hz beim Volleyprinzip schließt eine Chromawahrnehmung für Töne über dieser Schwelle aus. Da die zeitliche Analyse den Residualton beeinflusst, kann auch diesem ein Chroma zugeordnet werden. Im Gegensatz dazu beruht die Tonhelligkeit auf der Frequenzanalyse. Somit kann sowohl periodischen als auch nicht-periodischen Klängen eine Tonhelligkeit zugeordnet werden (vgl. Abschnitt 1.1.3).

Eine detaillierte Diskussion der Tonhöhenwahrnehmungsmodelle findet sich in Abschnitt 7.

¹⁶¹Hesse (2003b), Seite 130

¹⁶²Hesse (2003b), Seite 131

6 Mathematische und Neuronale Autokorrelation

6.1 Mathematische Autokorrelation

Um festzustellen, ob sich ein Signal nach einer gewissen Periode wiederholt, wird in der Signalverarbeitung üblicherweise eine Autokorrelationsfunktion verwendet. Etwaige Perioden können am Auftreten globaler Maxima erkannt werden. Gibt es nur *ein* globales Maximum, so ist die Funktion nicht periodisch. Ansonsten kann am Abstand der globalen Maxima die Periodenlänge abgelesen werden.

Es sei an dieser Stelle festgehalten, dass der Zusammenhang zwischen Autokorrelation und Fouriertransformation durch das Theorem von Wiener-Chintschin beschrieben wird. Dieses besagt im Wesentlichen, dass Autokorrelation und Fouriertransformation über die Leistungsdichte miteinander in Verbindung gebracht werden können. Ein Beweis dieser Aussage findet sich bspw. in Gnedenko, *Einführung in die Wahrscheinlichkeitstheorie*.¹⁶³

6.1.1 Definition der Autokorrelationsfunktion

Man betrachtet eine Funktion $f(t)$ aus dem Raum L^2 . Das bedeutet, dass die Funktion $|f|^2$ (Lebesgue) integrierbar und der Wert des Integrals eine reelle Zahl ist.

Dann kann die zugehörige Autokorrelationsfunktion $A_f(t)$ wie folgt definiert werden:

$$A_f(t) := \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(\tau) f(\tau + t) d\tau \quad t \in \mathbb{R} \quad (4)$$

Das bedeutet, dass man für einen festen Zeitpunkt t die Funktion mit der um t verzögerten Funktion multipliziert und die erhaltenen Werte in einem bestimmten Zeitraum aufsummiert bzw. integriert. Der Limes steht für den Grenzwert, wenn dieser Zeitraum sehr groß wird. $\frac{1}{2T}$ ist ein gewichtender Faktor.

¹⁶³Gnedenko (1991), Seite 307

6.1.2 Eigenschaften der Autokorrelationsfunktion

Die Autokorrelationsfunktion liefert für alle Zeitpunkte t einen nichtnegativen reellen Wert. Der Beweis folgt unmittelbar aus der Cauchy-Schwarz-Ungleichung,¹⁶⁴ die in allen Hilberträumen (Vektorraum mit Skalarprodukt) gilt.¹⁶⁵

$$0 \leq \left| \frac{1}{2T} \int_{-T}^T f(\tau) f(\tau + t) d\tau \right|^2 \quad (5)$$

$$\leq \left[\frac{1}{2T} \int_{-T}^T f^2(\tau) d\tau \right] \left[\frac{1}{2T} \int_{-T}^T f^2(\tau + t) d\tau \right] \xrightarrow{T \rightarrow \infty} A_f(0) A_f(0) \in \mathbb{R} \quad (6)$$

Aus dieser Ungleichungskette folgt:

$$0 \leq |A_f(t)| = \left| \frac{1}{2T} \int_{-T}^T f(\tau) f(\tau + t) d\tau \right| \leq A_f(0) \quad \text{für alle } t \in \mathbb{R}$$

Man erhält insgesamt:

$$|A_f(t)| \leq A_f(0) \quad \text{für alle } t \in \mathbb{R}$$

Das bedeutet, dass die Autokorrelationsfunktion für $t = 0$ ein globales Maximum annimmt.

Man kann zeigen, dass für periodische Funktionen $f(t)$ die Autokorrelationsfunktion $A_f(t)$ an allen Vielfachen der Periode T ebenfalls den Maximalwert annimmt. Dies betrifft alle Stellen mit $t = nT$ für $n \in \mathbb{N}$.

Man betrachte dazu die Funktion $A_f(nT)$, dann folgt die Behauptung unmittelbar durch Einsetzen der Definition auf Grund der Periodizität mit $f(\tau + nT) = f(\tau)$:¹⁶⁶

¹⁶⁴Diese Ungleichung besagt: $|\int f g|^2 \leq \int |f|^2 \int |g|^2$

¹⁶⁵Beweis geführt nach Papoulis (1962), Seite 245

¹⁶⁶Ebeling (2007), Seite 157

$$A_f(nT) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(\tau) f(\tau + nT) d\tau \quad (7)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(\tau) f(\tau) d\tau \quad (8)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(\tau) f(\tau + 0) d\tau \quad (9)$$

$$= A(0) \quad (10)$$

Auf diese Weise kann man mit Hilfe der zugehörigen Autokorrelationsfunktion die Periode der Funktion f finden, indem man die Stellen bestimmt, an denen ein globales Maximum angenommen wird und von allen diesen Stellen das Minimum auswählt. Gibt es nur *ein* globales Maximum, so ist die Funktion nicht periodisch.

6.2 Neuronale Autokorrelation

Bei der Suche nach den biologischen Mechanismen, die eine Erklärung für das Residualtonhören liefern könnten, wurde bald klar, dass weder die Wellenform noch die Einhüllende des Signals allein für die Residualtonwahrnehmung verantwortlich sein können. Hartmann formulierte diesen Erkenntnisprozess wie folgt:¹⁶⁷

„The problem with the peaks and valleys model is that these are waveform features that depend sensitively on the phases of the components. These phases change dramatically as one moves from place to place in a room, and yet we find that the pitch of a tone does not change as we walk around a room. [...] Therefore the envelope periodicity model is more attractive. However, we can create a problem for the envelope periodicity model by generating a complex tone with all its harmonics shifted by a common frequency difference [...]. [...] The envelope depends only on the spacing of the harmonics. Therefore, the envelope periodicity model predicts that the pitch should be unchanged by the frequency shift. The problem is that experiments show clearly that the pitch does change [...].“

¹⁶⁷Hartmann (1996), Seite 3498

Da weder die Wellenform noch die Einhüllende zufriedenstellende Erklärungen lieferten, wurde nach anderen Ansätzen gesucht.

Seit der ersten Hälfte des 20. Jahrhunderts gibt es die Theorie, dass in der Hörbahn des Menschen zur Analyse akustischer Signale auch eine Autokorrelation abläuft. Die daraus gewonnene Periodotopie (Abbild der Periodizitätsanalyse) stellt neben der Tonotopie (Abbild der Frequenzanalyse) die Grundlage für die akustische Einordnung von Klängen dar.

Die neuronale Autokorrelation entpuppte sich tatsächlich als Kandidat, um bisher durch die Frequenzanalyse nicht oder nur unzureichend erklärbare Phänomene wie das Residualtonhören begründen zu können.

Die biologischen Voraussetzungen dafür wurden inzwischen im menschlichen Gehirn u.a. von Langner nachgewiesen. Dabei stellte sich heraus, dass die Autokorrelationsfunktion aus den Modellen die all-order-ISI-Histogramme der Nervenimpulse beschreiben konnte.

Im folgenden Kapitel werden nun unterschiedliche Modelle, die mittels neuronaler Autokorrelation Aspekte der Tonhöhenwahrnehmung auf biologische, neurologische bzw. mathematische Weise erklären, präsentiert und verglichen.

7 Neuronale Modelle der Tonhöhenwahrnehmung

Wenige Jahre nach Schoutens Veröffentlichungen über das Residualtonhören (vgl. Abschnitt 4.3.1) schrieb Licklider den Artikel *Duplex Theory of Pitch Perception* (1951). Für diese Theorie gab es zum Zeitpunkt des Erscheinens noch keine neurologische Evidenz. Erst später konnten Wissenschaftler wie Langner zeigen, dass die biologischen Grundlagen dafür beim Menschen vorhanden sind (vgl. Abschnitt 7.1.2). Durch die Verbindung der Orts- und der Zeittheorie lieferte Lickliders Modell Erklärungen für Beobachtungen, die keine der beiden Theorien allein bieten konnte.

Physikalisch gesehen wurde die Tonhöhe schon längere Zeit als duale Erscheinung aufgefasst, einerseits durch die Anzahl der Schwingungen pro Sekunde (Frequenz f), andererseits durch die Periodendauer zwischen zwei Schwingungsmaxima (Periode T). Der Zusammenhang lässt sich in der Physik mit $f = \frac{1}{T}$ beschreiben.

Licklider ging nun davon aus, dass das menschliche Ohr beide Attribute zur Tonhöhenwahrnehmung verwende. Die Frequenzanalyse in der Cochlea galt seit Békésys Forschungsergebnissen über die Basilarmembran als gesichert. Hingegen blieb die Möglichkeit einer Autokorrelationsanalyse im Zeitbereich bis zu Lickliders Publikation weitgehend unbeachtet.

Mathematisch gesehen erklärt das Wiener-Chintschin-Theorem, wie Autokorrelation, Leistungsspektrum und Fouriertransformation zusammenhängen (vgl. Abschnitt 6.1.1). Es stellt sich nun die Frage, welche Informationen das Gehör durch eine Autokorrelationsanalyse, bei der die Periodendauern der Bestandteile eines Klages als Grundlage dienen, zusätzlich gewinnen kann.

Im folgenden Kapitel werden einige Tonhöhenwahrnehmungsmodelle mit Zeitbereichsanalyse bzw. Autokorrelation vorgestellt. Im Speziellen wird bei diesen Modellen die Fähigkeit zur Erklärung des Residualtonhörens untersucht.

7.1 Lickliders *Duplex Theory of Pitch Perception*

7.1.1 Formalisierung des neuronalen Autokorrelationsprozesses nach Licklider

Licklider veröffentlichte 1951 den Artikel *A Duplex Theory of Pitch Perception* und präsentierte darin ein Tonhöhenwahrnehmungsmodell, das auf Frequenz- und Periodizitätsanalyse beruhte. Zum Zeitpunkt der Veröffentlichung konkurrierten immer noch Zeit- und Frequenzanalysemodelle um die Erklärungshoheit der Tonhöhenwahrnehmung. Der Ansatz, diese beiden Modelle zu verbinden, war zum damaligen Zeitpunkt einzigartig und öffnete einen völlig neuen Anknüpfungspunkt.

Licklider¹⁶⁸ versah die abgewinkelte Cochlea mit einer x -Koordinate, die den Abstand eines Punktes vom basalen Ende der Schnecke beschrieb. Durch die Frequenzanalyse in der Cochlea wird die Schallwelle $f(t)$ in eine Funktion $F(x, t)$ umgewandelt, die die Amplitude der Frequenzkomponente x zur Zeit t beschreibt. Wobei die Amplitude der Auslenkung der Cochlea entspricht.

Die Neuronen des Gehörnervs, die an der Stelle x_0 der Cochlea ansetzen, leiten die Erregung $F(x_0, t)$ in Richtung Gehirn weiter. Genauer gesagt, entspricht jede Funktion $F_x(t) := F(x, t)$ in der Realität dem Integral über mehrere Neuronen. Dabei kann deren Feuerverhalten gemäß den biologischen Gegebenheiten jeweils mit 0 (nichtfeuernd) oder 1 (feuernd) beschrieben werden.

Im Folgenden wird eine schematische Übersicht der biologischen Abläufe der Lickliderschen Theorie präsentiert. Auch wenn im Ohr stets mehrere Neuronen zusammenwirken, wird aus Gründen der Übersichtlichkeit hier jeweils nur von einem einzelnen Neuron gesprochen.¹⁶⁹

1. Neuron A hängt direkt am Eingangssignal („input neuron“). Je nachdem feuert es oder nicht.

¹⁶⁸Licklider (1951), Seite 128

¹⁶⁹Licklider (1951), Seite 129

2. Neuronen B_1 bis B_k feuern mit einer Zeitverzögerung von $k\tau$, wobei τ eine sehr kurze Zeiteinheit beschreibt. In diesem Schritt wird ein verzögertes Signal gebildet.
3. Zu den Neuronen B_1 bis B_k gibt es weiterführend Neuronen C_1 bis C_k , wobei Neuron C_i mit $i \in 1, \dots, k$ genau dann feuert, wenn das Neuron A und das Neuron B_i beide gleichzeitig feuern. (Anmerkung: Im mathematischen Modell kann von Gleichzeitigkeit gesprochen werden, biologisch gesehen ist innerhalb eines sehr kurzen Zeitintervalls gemeint.) Man spricht dabei von Multiplikation im neurologischen Sinn.
4. Der nächste Schritt ist die Integration. Ein Neuron D_i summiert die empfangenen Impulse des Neurons C_i auf, bis diese einen gewissen Schwellenwert übersteigen. Dann wird ein Aktionspotential ausgelöst, das bis in das Gehirn weitergeleitet wird.

Dieser Autokorrelations-Vorgang passiert unmittelbar nach der Frequenzanalyse der Cochlea. Licklider zog auch die Möglichkeit in Betracht, dass das Neuron C_i die Rolle des Neurons D_i übernimmt, und somit gleichzeitig für Multiplikation und Integration verantwortlich ist.

Licklider über die biologischen Hintergründe zu seiner Autokorrelations-Theorie:¹⁷⁰

„A chain of neurons makes an excellent delay line. The spatial aspect of synaptic summation provides approximate multiplication. And the temporal aspect of synaptic summation gives us running integration. The theory postulates, therefore, that the lower centers of the auditory system perform the three operations.“

Licklider verknüpfte mit dieser Theorie erstmals zwei Aspekte eines Klanges, die Frequenz und die Periodizität und entwarf eine neurologische Anordnung, die diese Analysen durchführen konnte. Seine Idee inspirierte viele Wissenschaftler, ähnliche Ansätze zu entwickeln, auch wenn sich später herausstellte, dass Lickliders Modell modifiziert werden

¹⁷⁰Licklider (1951), Seite 129

müsse, wie Ebeling es formulierte:¹⁷¹

„Das Modell von Licklider entspricht aber nicht den realen Verhältnissen in neuronalen Systemen. Die angenommene Verzögerungskette ist unphysiologisch, weil einer Vielzahl von Neuronen allein die Aufgabe der Signalverzögerung zukommt. Es ist daher nicht verwunderlich, dass es in keinem von den bisher untersuchten Hörsystemen eine Evidenz dafür gibt. Wenn man also einen Autokorrelator im Hörsystem annimmt, so müssen die notwendigen Verzögerungen mit physiologisch anderen Mitteln erzielt werden.“

Tramo et al. führten an, dass bis heute keine entsprechenden biologischen Nachweise für Lickliders Modell gefunden werden konnten:¹⁷²

„Psychophysical data on the pitch of inharmonic tones strongly suggest a central temporal autocorrelation analysis rather than a modulation or envelope-based one. Licklider’s 1951 time-delay (TDNN) autocorrelation network could carry out the right time-place transformation, but no units with requisite fine comb-filter tunings have been found at brain stem levels. Modulation tunings are seen in many units in the auditory thalamus and cortex, [...] but they are generally very coarse [Anm.: rauh] compared to pitch discrimination thresholds.“

Zur Zeit von Licklider gab es noch keine biologischen Beweise für eine Autokorrelationsanalyse im menschlichen Ohr. Erst einige Jahrzehnte später gelang es unter anderem Langner Belege für ein Autokorrelations-Modell in Säugetieren und später im Menschen zu finden.

7.1.2 Biologische Evidenz für ein Autokorrelations-Modell nach Langner

Gerald Langner konnte mit seinen Untersuchungen Lickliders Autokorrelations-Theorie biologisch untermauern. Es gelang ihm zu zeigen, dass bei der Tonhöhenwahrnehmung sowohl die Frequenz als auch die Periodizität in den Stationen der neuronalen Verarbeitung vorhanden sind. Dies bestärkt die duale Auffassung eines Tones, der einerseits

¹⁷¹Ebeling (2009), Seite 77

¹⁷²Tramo u. a. (2005), Seite 163

durch Frequenz und andererseits durch Periodizität beschrieben werden kann:¹⁷³

„Etwas vereinfacht lässt sich also feststellen, dass der Klang eines Tones weitgehend, wenn auch nicht ausschließlich, von seinem Frequenzspektrum bestimmt wird, die Tonhöhe dagegen von der Periodizität seiner Wellenform. In diesem Sinne sind Frequenz und Tonhöhe tatsächlich nahezu unabhängige Größen. Sie [...] werden auch im Hörsystem weitgehend separat verarbeitet. Wie im Folgenden dargestellt, werden sie ihrer Unabhängigkeit entsprechend im auditorischen Mittelhirn und auch im Cortex räumlich orthogonal zueinander abgebildet.“

Langner hielt fest, dass es nur bei den tiefen Frequenzen eine eindeutige Zuordnung zwischen Frequenz und innerer Haarzelle gibt. Bei den höheren Frequenzen kann eine Haarzelle von einem ganzen Frequenzbereich angesprochen werden. Diese Erkenntnis trug zu einem besseren Verständnis des Residualtonphänomens („missing fundamental“) bei.¹⁷⁴

„Die begrenzte Frequenzauflösung der Cochlea hat also einen großen Vorteil. Haarzellen, die auf Grund ihrer Position in der Cochlea auf bestimmte Obertöne reagieren, übertragen wegen der Überlagerung dieser Obertöne mit benachbarten Obertönen zusätzlich zeitliche Information über die Periode des Grundtons in das zentrale Hörsystem. Aus diesem Grund kann die Überlagerung der Obertöne nach entsprechender Analyse dieselbe Tonhöhenwahrnehmung auslösen wie die Grundfrequenz allein („missing fundamental“).“

Für jeden Frequenzbereich wird dabei die Periodizität bestimmt und durch das Feuermuster des Hörnervs an die Nuclei Cochleares (CN, Teil des Hirnstammes) übertragen, wo die unterschiedlichen Zelltypen eine Korrelationsanalyse ermöglichen.

Zelltypen für die Autokorrelationsanalyse: Die Autokorrelationsanalyse wird in Langners Modell von mehreren Zelltypen durchgeführt, von denen jeder eine spezielle Aufgabe erfüllt.

¹⁷³Langner (2007), Seite 11

¹⁷⁴Langner (2007), Seite 13

Zunächst werden Zellen benötigt, die die Impulse möglichst unverzerrt weiterleiten, um das Eingangssignal aus beiden Ohren vergleichen zu können.

Dafür sind die Buschzellen („bushy cells“) in den beiden Nuclei cochleares verantwortlich. Sie sorgen dafür, dass in den Zellen der Nuclei olivares, wo erstmals in der Hörbahn akustische Informationen aus beiden Ohren zusammenfließen, eine zeitliche Analyse stattfinden kann. Die dortige Messung der Laufzeitdifferenzen von bis zu 10 μ s (Langner¹⁷⁵) bzw. 30 μ s (Hesse¹⁷⁶) ermöglicht eine äußerst präzise Bestimmung der Position der Schallquelle im Raum und auf diese Weise das Richtungshören.

Die Octopus-Zellen („on-Zellen“) in den Nuclei cochleares können Signale über weite Bereiche der Basilarmembran integrieren. Im Modell sind sie dafür verantwortlich, die anderen Neuronengruppen zeitlich abzustimmen. Sie arbeiten sehr breitbandig, im Gegensatz zu den von ihnen aktivierten Sternzellen („Chopper-Neuronen“). Deren ständige Oszillationen schaffen eine zeitliche Referenz, die für das absolute Gehör verantwortlich sein könnte, wie Langner vermutete.¹⁷⁷

Ebenso werden in den Nuclei cochleares (CN) die Spindelzellen, auch Integratorneuronen genannt, von den Octopus-Zellen zeitlich synchronisiert zur Integration angeregt. Im Vergleich zu den Sternzellen, die auch Oszillatorneuronen heißen, arbeiten sie in einem noch schmäleren Frequenzbereich.

Durch diese beiden Integrationen und die Informationsweiterleitung in die Colliculi inferiores (IC) kann dort eine neuronale Korrelation durchgeführt werden. Das bedeutet, dass die Informationen über Periode und Spektrum kombiniert und zu einer „räumlichen Repräsentation spektro-temporaler Information“ zusammengefügt werden.¹⁷⁸

Mittelhirn und Periodizität In Langners Korrelationsmodell findet die Koinzidenzdetektion in den Colliculi inferiores, einem Teil des Mittelhirns, statt.

¹⁷⁵Langner (2007), Seite 13

¹⁷⁶Hesse (2003b), Seite 129

¹⁷⁷Langner (2007), Seite 13

¹⁷⁸Langner (2007), Seite 14

Koinzidenzneuronen analysieren die Impulse der Stern- und Spindelzellen und reagieren auf zeitliche Übereinstimmungen der Signale. Dieser Vorgang entspricht einer Autokorrelation:

Das Triggerneuron ist für die Synchronisation der Messung verantwortlich. Ein Signal beschreibt das unverzögerte Eingangssignal, das von den Oszillatorneuronen geliefert wird. Das zweite Signal entspricht dem von den Integratorneuronen verzögerten Eingangssignal. Beide Signale gelangen zum Koinzidenzneuron, welches diese im nächsten Schritt auf Gleichheit untersucht.¹⁷⁹

Da es einer Neuronen-Gruppe nicht möglich ist, eine Autokorrelation mit variabler Verzögerungszeit durchzuführen, erfolgt der Vorgang an mehreren Orten gleichzeitig mit jeweils unterschiedlicher fester Verzögerungszeit. Fricke¹⁸⁰ spricht von einer „bank of parallel auto-correlators each of which shows fixed lag periods.“

Insgesamt liefern die neurologischen Abläufe somit nicht nur die Voraussetzung für eine tonotope Repräsentation des akustischen Klanges; erwiesenermaßen findet sich im IC zusätzlich auch eine periodotope Struktur.¹⁸¹

Cortex und Periodizität: Im Cortex befinden sich im Anfangsbereich ebenfalls eine tonotop organisierte und eine periodotope Struktur. Generell ist die Verarbeitung der akustischen Signale im Gehirn noch nicht restlos geklärt. Allerdings zeigt sich, dass „[...]die erste Achse im neuronalen Cortex die cochleäre Frequenzanalyse und damit die Frequenzkomposition oder Klanginformation eines Signals repräsentiert (Tonotopie), [während] sich auf der zweiten neuronalen Achse die Tonhöhe ab[bildet] (Periodotopie).“¹⁸²

Es existiert somit vom Beginn der akustischen Analyse im Innenohr bis zum Hörcortex stets die duale Repräsentation - Frequenz, Periodizität - in allen Verarbeitungsstationen. Beide Informationen stehen dem Gehirn bei der Tonhöhenwahrnehmung zur Verfügung.

¹⁷⁹Langner (2007), Seite 15

¹⁸⁰Fricke (2011), Seite 107

¹⁸¹Langner (2007), Seite 14

¹⁸²Langner (2007), Seite 18

Colliculi inferiores und Spiralstruktur in den VNLL: Durch die Korrelationsanalyse wird ein Neuron in den Colliculi inferiores nicht nur dann angeregt, wenn seine Frequenz der Grundfrequenz des Klanges entspricht, es reagiert auch auf die harmonischen Obertöne, bzw. auf die Subharmonischen ($\frac{f}{2}, \frac{f}{3}, \dots$).¹⁸³

Besonders bemerkenswert ist ferner, dass sich in den ventralen Nuclei lemnisci lateralis (VNLL), die für die Hemmung der Colliculi inferiores bei der Korrelationsanalyse verantwortlich sind,¹⁸⁴ eine Spiralstruktur findet, bei der eine Windung genau einer Oktave entspricht. Somit sind Töne im Oktavabstand übereinander geschichtet. Eine Ordnung, die an die Zweikomponententheorie der Tonhöhe erinnert (vgl. Abschnitt 1.1.3). Diese biologische Gegebenheit könnte erklären, wieso in Tonsystemen unterschiedlicher Musikkulturen die Oktave stets eine Sonderrolle einnimmt.¹⁸⁵

7.2 Ein stochastisches Modell der Tonhöhenwahrnehmung von Goldstein

Goldstein¹⁸⁶ präsentierte 1973 ein Tonhöhenwahrnehmungs-Modell, das auf einem Maximum-Likelihood-Schätzer beruht. Auch diese Theorie beinhaltet einen Mustererkennungsprozess.

Die zentrale Idee basiert darauf, dass die einzelnen Frequenzkomponenten als nebeneinander liegende harmonische Obertöne interpretiert werden und aus ihnen die zu Grunde liegende Grundfrequenz, welche die höchste Eintrittswahrscheinlichkeit aufweist, ermittelt wird.

Goldstein traf für sein Modell folgende drei Annahmen, wofür er empirische Belege von Helmholtz, Ritsma, Schouten, De Boer, etc. nannte:¹⁸⁷

¹⁸³Langner (2007), Seite 18

¹⁸⁴Langner (2007), Seite 13

¹⁸⁵Langner (2007), Seite 19–20

¹⁸⁶Goldstein (1973)

¹⁸⁷Goldstein (1973), Seite 1497 und 1509

1. Der zentrale Prozessor („central processor“) interpretiert die tonotopie Information stets als sinusförmige, direkt nebeneinander liegende, harmonische Obertöne.
2. Die Amplitude und Phase wird für die Analyse des Prozessors nicht berücksichtigt.
3. Die Frequenzinformationen können als unabhängige stochastische Signale interpretiert werden.

7.2.1 Mathematische Formulierung des Goldstein-Modells

Das Signal X_k aus der Frequenzanalyse in der Cochlea ist eine stochastische Größe und wird als normalverteilt angenommen. Der Mittelwert ist dabei die Frequenz f_k , die Standardabweichung σ_k wird als variabel angesehen und muss erst bestimmt werden. Es wird angenommen, dass die Standardabweichung lokal konstant ist und ihre Werte aus den früher von Houtsma und Goldstein gewonnenen empirischen Daten ermittelt werden können.¹⁸⁸

Der zentrale Prozessor („central processor“) schätzt nun den Residualton mit der Maximum-Likelihood-Methode. Das bedeutet, dass für die gegebenen Frequenzen die plausibelsten Positionen in der Obertonskala ermittelt werden und anschließend der zugehörige Grundton bestimmt wird. Plausibel bedeutet hier, dass die gewählten Parameter die Ergebnisse der ursprünglichen Frequenzanalyse am besten erklären.

Die Idee, einem Ton nicht eine eindeutige Frequenz zuzuordnen, sondern ihn als Wahrscheinlichkeitsdichtefunktion aufzufassen, versucht der musikalischen Unschärfe bei der Realisation eines Tones Rechnung zu tragen. Ähnliche Ansätze findet man bspw. bei Fricke (vgl. Abschnitt 7.3.2), Hartmann (vgl. Zitat in Abschnitt 1.2) und Ebeling (Töne als offene Intervalle betrachtet¹⁸⁹).

¹⁸⁸Houtsma u. Goldstein (1972) Seite 1499

¹⁸⁹Ebeling (1999), Seite 29

7.2.2 Der Residualton in Goldsteins Modell

Goldstein verwendete Schoutens und Ritsmas Daten um sein Modell zu testen. Dabei stellte er fest, dass seine Resultate — ähnlich wie bei Ritsma — einen dominanten Frequenzbereich bei der Residualtonwahrnehmung aufwiesen.

Die Frage, ob es biologische Hinweise dafür gibt, dass Neuronen einen Maximum-Likelihood-Schätzer ermitteln können bzw. wie dieser Prozess ablaufen könnte, wurde in Goldsteins Artikel nicht beantwortet.

7.3 Ein Autokorrelationsmodell von Impulsfolgen

7.3.1 Mathematische Formulierung des Autokorrelationsmodells

Der Musikwissenschaftler Martin Ebeling¹⁹⁰ formulierte 2007 ein Modell, das mit mathematischen Methoden die Vorgangsweise der Neuronen bei der Tonhöhenwahrnehmung beschreibt. In weiterer Folge leitete er daraus eine Konsonanztheorie ab.

Er baute dabei auf Cariani und Delguttess Arbeiten auf,¹⁹¹ die bei Messung festgestellt hatten, dass die Verteilung der all-order-ISI (interspike Intervalle) genau die neuronal codierte Information für die Tonhöhe beschreibt. Das bedeutet, dass die all-order-ISI-Histogramme äquivalent zur Autokorrelation des akustischen Reizes sind (vgl. Abschnitt 2.2.1).

Ebeling verwendete für sein Modell weiters die Ideen von Licklider (vgl. Abschnitt 7.1), der in seiner *Duplex Theory of Pitch Perception* bereits eine Autokorrelationsanalyse integriert hatte, sowie Resultate von Langner, der beim Menschen die Voraussetzungen für eine Autokorrelationsanalyse in den Colliculi inferiores nachweisen konnte. (vgl. Abschnitt 7.1.2)

Ebeling beschrieb den Ablauf der Koinzidenzdetektion bei der Autokorrelation wie folgt:¹⁹²

¹⁹⁰Ebeling (2007)

¹⁹¹Ebeling (2007), Seite 7

¹⁹²Ebeling (2008), Seite 2322

„At the beginning of stimulation, each coincidence neuron shows a response characteristic of a comb filter which makes the whole bank of coincidence circuits act as an autocorrelator for the modulation frequencies. However, after about 30 ms, inhibition triggered by the onset neuron converts the coincidence neuron to a bandpass filter [...] so that an unambiguous pitch discrimination is granted.“

Die Autokorrelationsfunktion eines Signals $f(t)$ lautet wie oben erwähnt (vgl. Abschnitt 6.1.1):

$$A_f(\tau) = \int f(t)f(t + \tau)dt$$

Ihr Maximum findet sich außer am Zeitpunkt $t = 0$ auch an allen Vielfachen der Periode T , also an den Zeitpunkten nT mit $n \in \mathbb{N}$.

Da der menschliche Hörbereich nach unten mit ca. 20 Hz begrenzt ist, genügt es, die Autokorrelation über ein Zeitintervall von 50 Millisekunden zu betrachten. Denn $\frac{1}{20}\text{sec} = 50\text{msec}$.

Es bleibt die Frage, welches Eingangssignal $f(t)$ für die Berechnung gewählt werden soll. Ebeling verwendete eine Rechtecksfunktion I_ϵ , die einer Wahrscheinlichkeitsdichtefunktion mit der Breite ϵ entspricht. Ein Sinuston der Frequenz $f_0 = \frac{1}{T_0}$ wird als eine periodische Abfolge dieser Rechtecke, mathematisch gesehen somit als Summe dieser Rechtecksfunktionen, dargestellt:

$$f(t) = \sum_n I_\epsilon(t - nT)$$

Dies hat den Zweck, dass kleine Unschärfen bei der Realisation des Klanges berücksichtigt werden. So entspricht ein Ton in der musikalischen Praxis stets nur ungefähr der am Notenpapier gedruckten Tonhöhen-Idealvorstellung ; ganz zu schweigen davon, dass unterschiedlichen Stimmungssystem leicht verschiedene Ton-Repräsentanten zu Grunde liegen.

Besteht ein Klang aus Sinustönen unterschiedlicher Frequenzen, so wird dieser als die Summe der Sinustöne dargestellt. Diese Interpretation eines Klangs als Summe der Frequenzkomponenten entspricht genau der Frequenzanalyse in der Cochlea. Dieses Signal geht somit in weiterer Folge in die neuronale Verarbeitung ein. Es sei an dieser Stelle angemerkt, dass zwischen den beiden Sinustönen auftretende nichtlineare Effekte wie Schwebung, Maskierung, etc. in diesem Modell nicht berücksichtigt werden. Ebeling rechtfertigte die Vernachlässigung wie folgt:¹⁹³

„However, first and foremost nonlinearity affects signal strength and not the time patterns. If the effects of nonlinearity are comparatively small, especially in the dynamic range of music, the consideration of a LTI system [Anm.: linear time-invariant system; lineares, zeitinvariantes System] is justified.“

Ebeling entwickelte das Modell anschließend noch weiter, um damit eine Konsonanztheorie zu begründen.¹⁹⁴

7.3.2 Fricke's Ergänzungen zu Unschärfe und Impulsform

Jobst Fricke führte in seinem Artikel *Neuronale Koinzidenz mit Unschärfe löst das Problem unbefriedigender Konsonanztheorien* in äußerst prägnanter Weise eine Zusammenfassung der Vorgänge in der Cochlea beim Hören aus. Zunächst betrachtete er den Fall, dass manche Teiltöne des Klangs innerhalb der kritischen Bandbreite liegen:¹⁹⁵

„Wenn an einer Stelle der BM einige stärkere Teiltöne (TT) enger benachbart liegen als eine kritische Bandbreite, das ist oberhalb c “ etwa eine Terz, pulsieren sie mit der Frequenz ihrer Differenz. Sie bilden sozusagen eine Schwebung; diese hat bei harmonischen TT die Frequenz des Residuums. Die Gruppe aus zwei oder mehr TT signalisiert dem Gehörorgan folgendes:

1. Der Ort der Reizung hat Bedeutung: es liegt eine diesem Ort entsprechende klangfarbliche Auswertung vor: Formantwirkung (tonotope Abbildung).

¹⁹³Ebeling (2008), Seite 2328

¹⁹⁴Ebeling (2007)

¹⁹⁵Fricke (2009), Seite 104

2. Grundtonfrequenz ist gegeben durch die Frequenz des Pulsierens der BM an dieser Stelle (periodotope Abbildung). Das heißt: Auswertung des Residuums findet an einer Stelle statt, die der klangfarblichen Auswertung höherer Frequenzen zugeordnet ist.“

Sobald zwei Obertonfrequenzen im hörbaren Bereich sehr nahe beieinander liegen, lösen sie eine Schwebung aus. In Folge beginnt die Basilarmembran in der Schwebungsfrequenz zu schwingen. Dadurch stehen sowohl periodotope als auch tonotope Informationen für die weitere Analyse zur Verfügung.

In einer typischen Residualtonsituation gibt es mindestens zwei harmonische Obertöne, die jedoch nicht zwingend innerhalb einer kritischen Bandbreite liegen müssen. In diesem Fall verändern sich die Abläufe auf der Basilarmembran leicht:¹⁹⁶

„Wenn an einer Stelle ein einzelner Teilton stärker hervortritt, dessen benachbarte TT mehr als eine Quarte Abstand von ihm haben, vibriert die Stelle mit seiner Frequenz und liegt auch an der dieser Frequenz zugeordneten Stelle. Sie signalisiert:

1. Aus der an diesem Ort vorliegenden Reizung folgt die klangfarbliche Wirkung dieser Komponente (tonotope Abbildung).
2. Die Bewegungsfrequenz an dieser Stelle ist die der Spektralkomponente (Der Reizort für die periodotope Abbildung ist der gleiche wie der für die tonotope).

Aus diesem Sachverhalt ist folgender Schluss zu ziehen:

1. Klangfarbe wird vermittelt durch den Ort auf der BM, und zwar unabhängig davon, mit welcher Frequenz diese Stelle vibriert.
2. Residuum kann an jeder Stelle der BM abgeleitet werden.
3. Verzögerungszeiten kommen auch durch die Ausbreitung der Welle auf der BM zustande.“

¹⁹⁶Fricke (2009), Seite 104

Auch in diesem Fall stehen den höheren Ebenen der Hörbahn Tonotopie und Periodotopie für die Bewertung des akustischen Reizes zur Verfügung.

In weiterer Folge beschäftigte Fricke sich eingehend mit der Unschärfe, die nötig ist, um Tonhöhenwahrnehmungsmodellen ein realistisches Antlitz zu geben. Der große Nachteil vieler Modelle ist, dass sie nur für exakt realisierte Frequenzen funktionieren. Dies entspricht jedoch keineswegs der musikalischen Praxis. Fricke verwendete den Begriff „Klangbreite“ für all diejenigen Realisationen eines Tones, die von Zuhörern als derselbe Ton wahrgenommen werden. Dieser Bereich kann im Extremfall bis zu einem Halbton umfassen (83 cents in Fricke's Untersuchungen¹⁹⁷) und hängt grundsätzlich vom Instrument, der Stimmung und den Versuchspersonen ab.

Intervalle wie Oktav oder Quint klingen für das menschliche Ohr auch dann konsonant, wenn eine leichte Verstimmung vorliegt. Fricke berichtete von Abweichungen bis zu 25 cents bzw. $\frac{1}{8}$ -Ton.¹⁹⁸ Bei Untersuchungen von Miskiewicz und Rogala¹⁹⁹ stellte sich sogar heraus, dass Hörer Intervalle, die nicht exakt den Frequenzverhältnissen entsprachen, bevorzugten. So wurde etwa eine temperierte Quint als weniger rauh empfunden, als eine reine Quint mit einem Frequenzverhältnis von 3:2.

Auch auf der Ebene der neuronalen Verarbeitung spielt die Unschärfe auf Grund biologischer Grenzen eine große Rolle, die in vielen Modellen zu wenig berücksichtigt wird:²⁰⁰ „Diese Unschärfe resultiert einerseits schon aus der logischen Konsequenz, dass neuronale Schaltprozesse nicht mit elektronischen Simulationen gleichzusetzen sind, sondern reale Vorgänge des Ionenaustauschs in Nervenzellen sind. Deshalb lieferten Modellrechnungen mit Zeitreihenanalysen auf der Basis der Autokorrelation bisher unbefriedigende Ergebnisse, die die Erfahrungen mit der praktischen Musikausübung nicht abbilden konnten.“

Aus der Existenz der neurologischen Unschärfe folgt unmittelbar die Frage, wie die Impulsform für die Autokorrelation gewählt werden soll, um die realen Gegebenheiten

¹⁹⁷Fricke (1988), Seite 67

¹⁹⁸Fricke (2009), Seite 96

¹⁹⁹Miskiewicz u. Rogala (2003), 2003

²⁰⁰Fricke (2009), Seite 99

möglichst gut widerzuspiegeln. Sehr schmale Impulse wie z.B. Dirac-Impulse würden eine sehr präzise neuronale Verarbeitung beschreiben, während sehr breite Impulse für eine große Unschärfe stünden. Fricke schilderte die Vor- und Nachteile der unterschiedlichen Zugänge wie folgt:²⁰¹

„Sinusförmige Komponenten des Schalls $f(t) = \sin \omega t$, aus denen der Schall nach Fourier [Anm.: vgl. Formel (1)] zusammengesetzt ist, ergeben nach dieser Formel [Anm.: vgl. Formel (4)] nur schwach von einander abgesetzte Maxima und damit eine schlechte Trennung der einzelnen Komponenten. Setzt man in die Autokorrelationsanalyse anstelle der Sinuskomponenten Nadelimpulse oder Dirac-Stöße ein, wird die Trennung so scharf, dass sie der Wirklichkeit neuronaler Verarbeitung nicht mehr entspricht. Sie liefert diskrete Frequenzpunkte bei genauer Koinzidenz, ohne Übergänge und Zwischenwerte.“

Eine Variation der Impulsform führt somit zu großen Unterschieden in der Modellierung und kann die Praxistauglichkeit eines Autokorrelations-Modelles stark beeinflussen.

Ein anderer Weg wäre, anstelle der Modifikation der Impulsfolge ein Zeitfenster für die Detektion der Übereinstimmung einzuführen. Fricke schlug für die Impulsbreite eine Dauer zwischen 0,005 msec und 1 msec vor. Stattdessen könnte auch ein Zeitfenster von etwa 0,8 msec verwendet werden, womit zusätzlich die Verzögerung durch neuronale Reaktionszeiten in der Modellbildung inkludiert ist.²⁰²

Fricke beschäftigte sich auch mit den neurologischen Abläufen des Autokorrelationsvorganges. Wie bereits erwähnt, wird für die Autokorrelation die Verzögerung des Eingangssignals benötigt. Fricke befürwortete eine neuronale Verzögerungskette und sah in ihr das Potential, die Residualtonwahrnehmung hinreichend zu erklären:²⁰³

„Die Kombination von Verzögerungszeiten, die man durch Umwertschaltungen und Verzögerungselemente erreichen kann, mit Und-Gattern, in denen die Gleichzeitigkeit festgestellt wird, bietet sich also grundsätzlich an, um Zeitunterschiede und Zeitstrecken zu messen. Bei Tönen besteht so die Möglichkeit, die Dauer ihrer Perioden T und damit

²⁰¹Fricke (2009), Seite 109

²⁰²Fricke (2009), Seite 115

²⁰³Fricke (2009), Seite 110

(indirekt) ihre Tonhöhe zu ermitteln. Würde dieses Verfahren zur Messung der Frequenz $f = \frac{1}{T}$ eingesetzt, würde gleichzeitig „the case of the missing fundamental“ [...] gelöst. Denn die Periode findet sich in jeder Gruppe benachbarter Obertöne wieder, die die Basilarmembran in Bewegung setzen.“

7.4 Residualton & Neuronale Autokorrelation

In den vorhergehenden Abschnitten wurden einige Tonhöhenwahrnehmungsmodelle, die auf Autokorrelation beruhen, untersucht. In diesem Abschnitt soll nun zusammengefasst werden, inwiefern die neuronale Autokorrelation das Residualtonhören erklären kann.

Die neuronale Autokorrelation ermöglicht es, mathematisch zu zeigen, dass ein im Frequenzspektrum fehlender Grundton keine Auswirkungen auf die Autokorrelationsfunktion hat. Das bedeutet, dass für die weiterverarbeitenden Gehirninstanzen kein Unterschied zwischen einer Autokorrelation des Signals mit Grundton und einer Autokorrelation des Signals ohne Grundton erkennbar ist.

Tramo et al. stellten dies in dem Artikel *Neurophysiology and Neuroanatomy of Pitch Perception: Auditory Cortex* ausführlich dar.²⁰⁴ Sie verglichen darin die Autokorrelationsfunktion eines Sinussignals, eines harmonischen Spektrums mit Grundton und eines harmonischen Spektrums ohne Grundton miteinander. Zwar unterschieden sich diese drei Autokorrelationsfunktionen grundsätzlich, doch in allen drei Fällen befand sich das erste Maximum der Autokorrelationsfunktion — ausgenommen das Maximum im Ursprung — genau an der Stelle der Periode der Grundfrequenz. Somit ergab die Autokorrelation für alle drei Fälle die Tonhöhenwahrnehmung des Grundtons, egal ob dieser vorhanden war oder nicht.

Somit konnte nachgewiesen werden, dass das über viele Jahrzehnte hindurch ungeklärte Residualtonphänomen mittels neuronaler Autokorrelation begründet werden kann. Besonders erfreulich ist, dass diese Theorie sich auch in der Praxis widerspiegelt. Dies

²⁰⁴Tramo u. a. (2005), Seite 150, fig.1

zeigten die Untersuchungen von Wissenschaftern wie Langer bzw. Tramo et al., die biologische und neurologische Evidenz für den Ablauf einer Autokorrelation beim Menschen finden konnten.

8 Fazit

Die jahrhundertealte Frage nach der Tonhöhenwahrnehmung beim Menschen kann immer noch nicht als vollständig beantwortet angesehen werden. Jedoch wurden speziell in den letzten hundert Jahren auf dem Gebiet große Fortschritte erzielt. Der Aufbau des menschlichen Ohres wurde ausführlich dokumentiert und die mechanischen Vorgänge im Innenohr geklärt. Auch die Frequenzanalyse in der Cochlea gilt mittlerweile als gesichert. Die neurologischen Abläufe beim Hören konnten hingegen noch nicht hinreichend erforscht werden. Jedoch stellt die biologische Entdeckung der Zeitbereichsanalyse einen erfolgsversprechenden Ansatz dar.

Somit stehen dem menschlichen Gehörsinn zwei unterschiedliche Aufbereitungen der akustischen Information für die Bewertung des Gehörten zu Verfügung: einerseits die Aufschlüsselung der Frequenzen und andererseits die Periodizitätsabstände. Das Zusammenwirken dieser beiden Analysemethoden ermöglicht es, die verschiedenen charakteristischen Eigenschaften und Phänomene des Hörsinns zu begründen. Zahlreiche neue Modelle zur Tonhöhenwahrnehmung wurden in den letzten Jahren entworfen und mit Daten aus Versuchsreihen auf ihre Gültigkeit überprüft. Die meisten verwenden neuronale Autokorrelation für eine Periodizitätsanalyse in den unteren Ebenen der Hörbahn. Vereinzelt finden sich auch Kritiker wie Hartmann,²⁰⁵ welcher bemängelt, dass Autokorrelationsmodelle bei tiefen Frequenzen versagen würden. Für die Detektion der Autokorrelation eines 100 Hz Sinustones müsste eine Verzögerung von 20 msec passieren. Doch bisher wären keine derart langen Verzögerungsketten im auditoren System bekannt.

Erfolgsversprechend ist indessen die Tatsache, dass mit der neuronalen Autokorrelation erstmals ein Instrument zur Erklärung des Residualtonhörens zur Verfügung steht. Auch wenn heutige Tonhöhenwahrnehmungsmodelle in mancherlei Hinsicht noch verbesserungswürdig erscheinen, so stellt die geglückte Verbindung von Frequenz- und Periodizitätsanalyse einen wichtigen Schritt für ein tieferes Verständnis des menschlichen Hörens dar.

²⁰⁵Hartmann (1996), Seite 3498

9 Ausblick

In dieser Arbeit wurde die neuronale Autokorrelation im Kontext der Tonhöhenwahrnehmung und speziell ihr Zusammenhang mit dem Residualtonhören untersucht. An dieser Stelle sei angeführt, dass durch den Einbau der Autokorrelation auch in anderen Fragestellungen wesentliche Fortschritte erzielt wurden. Ebeling²⁰⁶ und Tramo et al.²⁰⁷ verwendeten diesen Ansatz etwa, um ein funktionables Konsonanz/Dissonanz-Modell zu beschreiben.

Auf diese Weise erscheint es erstmals möglich, eine biologisch begründete Ursache für musikästhetische Empfindungen anzugeben.

Im Bereich des menschlichen Gehirns sind bekanntlich noch viele Fragen zur Tonhöhenwahrnehmung ungeklärt. Speziell die Funktionsweise des auditorischen Cortex entzieht sich noch größtenteils unserem Verständnis, ebenso wie die unterschiedlichen Funktionsweisen der Gehirnhälften. So gibt es beim Residualtonhören offenbar Unterschiede zwischen rechter und linker Hemisphäre. Paquette et. al. stellten in ihren Untersuchungen fest, dass bei den rechtshändigen Probanden das linke Ohr bzw. die rechte Hemisphäre einen Residualton rascher erkannte:²⁰⁸

„These results are in agreement with the hypothesis that the right hemisphere is dominant for abstract pitch computation while both hemispheres are capable of pitch perception when f_0 [Anm.: der Grundton] is present.“

Neben den psychoakustischen Hör-Experimenten gibt es Wissenschaftler, die versuchen, Teilprozesse der Tonhöhenwahrnehmung als unabhängig voneinander zu identifizieren. Dieser Vorgang wird in der Neurologie als Dissoziation bezeichnet. Tramo et al. fassten in einer 2005 veröffentlichten Publikation die Ergebnisse aus unterschiedlichen Studien zum Thema Tonhöhen- und Residualtonwahrnehmung bei Schäden am auditorischen Cortex wie folgt zusammen:²⁰⁹

²⁰⁶Ebeling (2007)

²⁰⁷Tramo u. a. (2001)

²⁰⁸Paquette u. a. (1996), Seite 155

²⁰⁹Tramo u. a. (2005), Seite 160

- „(1) frequency-selective neurons in A1 [Anm.: primäres auditorisches Gebiet des auditorischen Cortex] and adjacent areas in right and left TG [Anm.: Heschl's gyrus; Teil der Hörrinde] and STG [Anm.: Gyrus temporalis superior; Teil des Temporallappens] play critical, additive roles in our ability to detect a small change in pitch between two successive pure tones or harmonic tones;
- (2) right-sided neurons play a critical role in our ability to perceive the direction of a pitch change between two successive pure tones or harmonic tones;
- (3) left-sided neurons play an additive role in our ability to perceive the direction of a pitch change between two successive pure tones or harmonic tones;
- (4) right auditory cortex mechanisms mediating perception of F0 pitch versus missing-F0 pitch are neurologically dissociable; and
- (5) right auditory cortex mechanisms mediating spectral versus temporal processing of harmonic tones are neurologically dissociable.“

Wie aus diesen Zeilen hervorgeht, sind die genauen Abläufe noch unbekannt. Vermutlich bedarf es neuer technischer Entwicklungen, um weitgehende Klarheit über das Funktionieren des menschlichen Gehörsinns zu erlangen. Es bleibt zu hoffen, dass technische Innovationen im Bereich der Medizin bald Antworten auf die noch ungeklärten Fragen geben können. So ermöglicht etwa die hohe Rechenleistung heutiger Computersysteme die strömungsmechanische Modellierung komplexer Vorgänge, wie sie etwa in der Cochlea ablaufen.²¹⁰

Im Gegensatz dazu sind auf neuronaler Ebene und speziell im menschlichen Gehirn noch viele Fragen offen. So wirken die den Artikel abschließenden Worte von Tramo et al. etwas ratlos:²¹¹

„Thus, an enigma persists concerning the nature of the neural codes and computations that support these universal and highly reliable perceptual invariances and equivalence classes at the cortical level. Whatever the code, it must behave in a highly reliable way under a huge variety of environmental conditions and for stimuli that are almost ne-

²¹⁰Kanis u. de Boer (1996)

²¹¹Tramo u. a. (2005), Seite 170

ver present in the natural world. For the various reasons outlined here, we believe that the cortical representation of pitch is not likely to be based on idiosyncratic, special purpose assemblages of neural elements and interconnections. The neural representations and computations that subserve pitch perception may involve elegant and powerful information-processing principles and mechanisms that still elude us.“

Bedenkt man die großen Fortschritte, die seit Helmholtz in den letzten hundertfünfzig Jahren auf dem Gebiet der Tonhöhenwahrnehmung erzielt wurden, so kann man gespannt und optimistisch in die Zukunft blicken und sich auf baldige Antworten auf einige dieser offenen Fragen freuen.

10 Literatur

- [ASA 1960] *American Standard Acoustical Terminology*. American Standards Association, 1960
- [lex 1996] *Das neue Lexikon der Musik*. J.B. Metzler, 1996
- [Barkowsky 1996] BARKOWSKY, Johannes: *Das Fourier-Theorem in musikalischer Akustik und Tonpsychologie*. Bd. 8 der Reihe *Schriften zur Musikpsychologie und Musikästhetik* (Hrsg. Helga de la Motte-Haber). Peter Lang GmbH, 1996
- [von Békésy 1961] BÉKÉSY, Georg von: *Concerning the Pleasures of Observing, and the Mechanics of the Inner Ear (Nobel Lecture, 1961)*. http://www.nobelprize.org/nobel_prizes/medicine/laureates/1961/bekesy-lecture.html. Version: 1961, Abruf: 30.10.2012
- [von Békésy 1970] BÉKÉSY, Georg von: *Physiologie der Sinneshemmung*. München : Wilhelm Goldmann Verlag, 1970
- [Benninghoff u. Drenckhahn 2004] BENNINGHOFF, Alfred (Hrsg.) ; DRENCKHAHN, Detlev (Hrsg.): *Anatomie*. Bd. 2: *Anatomie. Makroskopische Anatomie, Histologie, Embryologie, Zellbiologie*. 16. Auflage. Elsevier, 2004
- [Bizley u. Walker 2010] BIZLEY, Jennifer K. ; WALKER, Kerry M. M.: Sensitivity and Selectivity of Neurons in Auditory Cortex to the Pitch, Timbre, and Location of Sounds. In: *The Neuroscientist* 16 (2010), Nr. 4, S. 453–469
- [Dostrovsky u. Cannon 1987] DOSTROVSKY, Sigalia ; CANNON, John T.: Entstehung der musikalischen Akustik (1600-1750). In: ZAMINER, Frieder (Hrsg.): *Geschichte der Musiktheorie Bd. 6: Hören, Messen und Rechnen in der frühen Neuzeit*. Darmstadt : Wissenschaftliche Buchgesellschaft, 1987, S. 7–80
- [Ebeling 1999] EBELING, Martin: *Tonhöhe physikalisch – musikalisch – psychologisch – mathematisch*. Bd. 2 der Reihe *Systemische Musikwissenschaft* (Hrsg. Jobst P. Fricke). Peter Lang GmbH, 1999

- [Ebeling 2007] EBELING, Martin: *Verschmelzung und neuronale Autokorrelation als Grundlage einer Konsonanztheorie*. Frankfurt am Main : Peter Lang GmbH, 2007
- [Ebeling 2008] EBELING, Martin: Neuronal periodicity detection as a basis for the perception of consonance: a mathematical model of tonal fusion. In: *The Journal of the Acoustical Society of America* 124 (2008), Oktober, Nr. 4, 2320–2329. <http://dx.doi.org/10.1121/1.2968688>. – DOI 10.1121/1.2968688. – ISSN 1520–8524
- [Ebeling 2009] EBELING, Martin: Zum Wesen der Konsonanz: Neuronale Koinzidenz, Verschmelzung und Rauigkeit. In: AUHAGEN, Wolfgang (Hrsg.): *Musikalisches Gedächtnis und musikalisches Lernen*. Hogrefe, 2009, S. 71–93
- [Endl u. Luh 2008] ENDL, Kurt ; LUH, Wolfgang: *Analysis II Eine integrierte Darstellung*. 8. Auflage. Wiesbaden : AULA-Verlag, 2008
- [Fastl u. Zwicker 2007] FASTL, Hugo ; ZWICKER, Eberhard: *Psychoacoustics Facts and Models*. 3. Auflage. Springer-Verlag, 2007
- [Fletcher 1924] FLETCHER, Harvey: The Physical Criterion for Determining the Pitch of a Musical Tone. In: *Physical Review* 23 (1924), S. 427–437
- [Fricke 1988] FRICKE, Jobst P.: Klangbreite und Tonempfindung. Bedingungen kategorialer Wahrnehmung aufgrund experimenteller Untersuchung der Intonation. In: BEHNE, K.-E. (Hrsg.) ; KLEINEN, G. (Hrsg.) ; MOTTE-HABER, H. de l. (Hrsg.): *Musikpsychologie, Empirische Forschungen – Ästhetische Experimente* Bd. 5. Wilhelmshaven : Florian Noetzel Verlag, 1988
- [Fricke 2009] FRICKE, Jobst P.: Neuronale Koinzidenz mit Unschärfe löst das Problem unbefriedigender Konsonanztheorien. In: AUHAGEN, Wolfgang (Hrsg.) ; BULLERJAHN, C. (Hrsg.) ; HÖGE, H. (Hrsg.): *Musikalisches Gedächtnis und musikalisches Lernen*. Hogrefe, 2009, S. 94–122
- [Fricke 2011] FRICKE, Jobst P.: A Statistical Theory of Musical Consonance is Proved in

- Praxis. In: SCHMIDHOFER, August (Hrsg.) ; JENA, Stefan (Hrsg.): *Klangfarbe*. Peter Lang GmbH, 2011, S. 99–112
- [Gelfand 2009] GELFAND, Stanley A.: *Essentials of Audiology*. Thieme, 2009
- [Gnedenko 1991] GNEDENKO, Boris W. ; ROSSBERG, Hans-Joachim (Hrsg.): *Einführung in die Wahrscheinlichkeitstheorie*. Berlin : Akademie Verlag GmbH, 1991
- [Goldstein 1973] GOLDSTEIN, Julius L.: An optimum processor theory for the central formation of the pitch of complex tones. In: *The Journal of the Acoustical Society of America* 54 (1973), Dezember, Nr. 6, 1496–1516. <http://dx.doi.org/10.1121/1.1914448>. – DOI 10.1121/1.1914448
- [Greenwood 1990] GREENWOOD, Donald D.: A cochlear frequency-position function for several species – 29 years later. In: *The Journal of the Acoustical Society of America* 87 (1990), Juni, Nr. 6, S. 2592–2605
- [Hällström 1832] HÄLLSTRÖM, Gustav G.: Von den Combinationstönen. In: *Annalen der Physik und Chemie* (1832), S. 438–466
- [Hartmann 1996] HARTMANN, William M.: Pitch, periodicity, and auditory organization. In: *The Journal of the Acoustical Society of America* 100 (1996), Dezember, Nr. 6, 3491–3502. <http://dx.doi.org/10.1121/1.417248>. – DOI 10.1121/1.417248
- [Haynes 1995] HAYNES, Bruce: Pitch. In: SADIE, Stanley (Hrsg.): *The New Grove Dictionary of Music and Musicians* Bd. 14. Osaka - Player piano. Macmillan Publishers Limited, 1995
- [Heffner u. Whitfield 1976] HEFFNER, Henry ; WHITFIELD, I. C.: Perception of the missing fundamental by cats. In: *The Journal of the Acoustical Society of America* 59 (1976), April, Nr. 4, 915–919. <http://dx.doi.org/10.1121/1.380951>. – DOI 10.1121/1.380951

- [von Helmholtz 1971] HELMHOLTZ, Hermann von ; KRAFT, Fritz (Hrsg.): *Über die physiologischen Ursachen der musikalischen Harmonien*. München : Kindler Verlag GmbH, 1971 [1857]
- [von Helmholtz 2003] HELMHOLTZ, Hermann von ; BRÜNING, Jochen (Hrsg.): *Gesammelte Schriften*. Bd. II: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Hildesheim : Georg Olms Verlag AG, 2003 [1863]
- [Hesse 2003a] HESSE, Horst-Peter: Hatte Lipps doch recht? Tonverwandtschaft und Tonverschmelzung im Lichte der heutigen Gehörpsychologie. In: NIEMÖLLER, Klaus W. (Hrsg.) ; GÄTJEN, Bram (Hrsg.): *Perspektiven und Methoden einer Systemischen Musikwissenschaft*. Peter Lang GmbH, 2003
- [Hesse 2003b] HESSE, Horst-Peter: *Musik und Emotion*. Wien : Springer-Verlag, 2003
- [Hoogland 1953] HOOGLAND, Gerrit A.: *The Missing Fundamental*. 1953
- [Houtgast 1976] HOUTGAST, T.: Subharmonic pitches of a pure tone at low S/N ratio. In: *The Journal of the Acoustical Society of America* 60 (1976), August, Nr. 2, 405–409. <http://dx.doi.org/10.1121/1.381096>. – DOI 10.1121/1.381096
- [Houtsma u. Goldstein 1972] HOUTSMA, A. J. M. ; GOLDSTEIN, J. L.: The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition. In: *The Journal of the Acoustical Society of America* 51 (1972), Februar, Nr. 2B, 520–529. <http://dx.doi.org/10.1121/1.1912873>. – DOI 10.1121/1.1912873
- [Hunt 1978] HUNT, Frederick V.: *Origins in Acoustics*. New Haven and London : Yale University, 1978
- [Kandel u. a. 1991] KANDEL, Eric R. (Hrsg.) ; SCHWARTZ, James H. (Hrsg.) ; JESSELL, Thomas M. (Hrsg.): *Principles of Neural Science*. 3. Elsevier, 1991

- [Kanis u. de Boer 1996] KANIS, Luc J. ; BOER, Egbert de: Comparing frequency-domain with time-domain solutions for a locally active nonlinear model of the cochlea. In: *The Journal of the Acoustical Society of America* 100 (1996), Oktober, Nr. 4, S. 2543–2546
- [Langner 2007] LANGNER, Gerald: Die zeitliche Verarbeitung periodischer Signale im Hörsystem: Neuronale Repräsentation von Tonhöhe, Klang, Harmonizität. In: *Zeitschrift für Audiologie* 46 (2007), Nr. 1, S. 8–27
- [Licklider 1951] LICKLIDER, J. C. R.: A Duplex Theory of Pitch Perception. In: *Experientia* VII (1951), S. 128–133
- [Meyer-Eppler u. a. 1959] MEYER-EPPLER, W. ; SENDHOFF, H. ; RUPPRATH, R.: Residualton und Formantton. In: *Gravesaner Blätter* IV (1959), Nr. XIV, S. 70–83
- [Miskiewicz u. Rogala 2003] MISKIEWICZ, Andrzej ; ROGALA, Tomira: Roughness and Dissonance of musical Dyads. In: KOPIEZ, R (Hrsg.) ; LEHMANN, A. C. (Hrsg.) ; WOLTER, I. (Hrsg.) ; WOLF, C. (Hrsg.) ; Hochschule für Musik, Theater und Medien (Veranst.): *Proceedings of the 5th Triennial ESCOM Conference*. Hannover, September 2003
- [Nelken u. Ahissar 2006] In: NELKEN, Israel ; AHISSAR, Merav: *High-level and Low-level Processing in the Auditory System: The Role of Primary Auditory Cortex*. IOS Press, 2006
- [Ohm 1839] OHM, Georg S.: Bemerkungen über Combinationstöne und Stöße. In: *Annalen der Physik und Chemie* XLVII (1839), S. 463–466
- [Ohm 1843] OHM, Georg S.: Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen. In: *Annalen der Physik und Chemie* LIX (1843), Nr. 8, S. 513–565
- [Papoulis 1962] PAPOULIS, Athanasios: *The Fourier Integral and its Applications*. McGraw-Hill Book Company, Inc., 1962 (McGraw-Hill Electronic Sciences Series)

- [Paquette u. a. 1996] PAQUETTE, Claude ; BOURASSA, Michelle ; PERETZ, Isabelle: Left ear advantage in pitch perception of complex tones without energy at the fundamental frequency. In: *Neuropsychology* 34 (1996), Nr. 2, S. 153–157
- [Rieländer 1982] RIELÄNDER, Michael M. (Hrsg.): *Reallexikon der Akustik*. Frankfurt am Main : Verlag Erwin Bochinsky, 1982
- [Ritsma 1962] RITSMA, Roelof J.: Existence Region of the Tonal Residue. I. In: *The Journal of the Acoustical Society of America* 34 (1962), September, Nr. 9, S. 1224–1229
- [Ritsma 1963] RITSMA, Roelof J.: Existence Region of the Tonal Residue II. In: *The Journal of the Acoustical Society of America* 35 (1963), August, Nr. 8, S. 1241–1245
- [Roeckelein 1998] ROECKELEIN, Jon E.: *Dictionary of Theories, Laws, and Concepts in Psychology*. Greenwood Press, 1998
- [Schneider u. Wengenroth 2009] SCHNEIDER, Peter ; WENGENROTH, Martina: The Neural Basis of Individual Holistic and Spectral Sound Perception. In: *Contemporary Music Review* 28 (2009), Juni, Nr. 3, 315–328. <http://dx.doi.org/10.1080/07494460903404402>. – DOI 10.1080/07494460903404402
- [Schouten 1938] SCHOUTEN, Jan F.: *The perception of subjective tones*. Koninklijke Nederlandsche Akademie van Wetenschappen, 1938
- [Schouten 1940] SCHOUTEN, Jan F.: The Perception of Pitch. In: *Philips Technical Review* (1940), S. 286–294
- [Schouten 1970] SCHOUTEN, Jan F.: The Residue Revisited. In: PLOMP, Reinier (Hrsg.): *Frequency analysis and periodicity detection in hearing*, 1970, S. 41–58
- [Schouten u. a. 1962] SCHOUTEN, Jan F. ; RITSMA, Roelof J. ; CARDOZO, B. L.: Pitch of the Residue. In: *The Journal of the Acoustical Society of America* 34 (1962), September, Nr. 9B, 1418–1424. <http://dx.doi.org/10.1121/1.1918360>. – DOI 10.1121/1.1918360

- [Seebeck 1841] SEEBECK, August: Beobachtungen über einige Bedingungen der Entstehung von Tönen. In: *Annalen der Physik und Chemie* LIII (1841), Nr. 7, S. 417–436
- [Seebeck 1843] SEEBECK, August: Ueber die Sirene. In: *Annalen der Physik und Chemie* LX (1843), Nr. 12, S. 449–481
- [Smith 1960] SMITH, Charles S.: *Leonhard Euler's Tentamen Novae Theoriae Musicae: A Translation and Commentary*. 1960
- [Terhardt 1976] TERHARDT, Ernst: Ein psychoakustisches Konzept der Musikalischen Konsonanz. In: *Acustica* 36 (1976/77), S. 121–137
- [Terhardt 1982] TERHARDT, Ernst: Die psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmische Bestimmung. In: DAHLHAUS, Carl (Hrsg.) ; KRAUSE, M. (Hrsg.): *Tiefenstruktur der Musik*. Technische Universität Berlin, 1982, S. 23–50
- [Tramo u. a. 2001] TRAMO, Mark J. ; CARIANI, Peter A. ; DELGUTTE, Bertrand ; BRAIDA, Louis D.: Neurobiological Foundations for the Theory of Harmony in Western Tonal Music. In: *Annals of the New York Academy of Sciences* 930 (2001), S. 92–116
- [Tramo u. a. 2005] TRAMO, Mark J. ; CARIANI, Peter A. ; KOH, Christine K. ; MAKRIS, Nikos ; BRAIDA, Louis D.: Neurophysiology and Neuroanatomy of Pitch Perception: Auditory Cortex. In: *Annals of the New York Academy of Sciences* 1060 (2005), S. 148–174
- [Turner 1977] TURNER, R. S.: The Ohm-Seebeck Dispute, Hermann von Helmholtz, and the Origins of Physiological Acoustics. In: *The British Journal for the History of Science* 10 (1977), March, Nr. 1, S. 1–24
- [Walliser 1969a] WALLISER, Konrad: Zur Unterschiedsschwelle der Periodentonhöhe. In: *Acustica* 21 (1969), S. 329–336
- [Walliser 1969b] WALLISER, Konrad: Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe. In: *Acustica* 21 (1969), S. 319–328

- [Wever u. Bray 1930] WEVER, Ernest G. ; BRAY, Charles W.: Action Currents in the Auditory Nerve in Response to Acoustical Stimulation. In: *Proceedings of the National Academy of Sciences of the United States of America* 16 (1930), Nr. 5, S. 344–350
- [Whitfield 1980] WHITFIELD, I. C.: Auditory cortex and the pitch of complex tones. In: *The Journal of the Acoustical Society of America* 67 (1980), Februar, Nr. 2, 644–647.
<http://dx.doi.org/10.1121/1.383889>. – DOI 10.1121/1.383889
- [Yost 1994] YOST, William A.: *Fundamentals of Hearing: an Introduction*. 3. Auflage. Academic Press, Inc., 1994

11 Abstracts in Deutsch und Englisch

11.1 Abstract in Deutsch

Das Ziel dieser Diplomarbeit war ein Vergleich der historischen und aktuellen Erkenntnisse in dem psychoakustischen Forschungsgebiet der Tonhöhenwahrnehmung. Zu Beginn wird ein historischer Überblick über die Ergebnisse von Physikern wie Seebeck, Ohm und Helmholtz gegeben und deren Einfluss auf die zukünftige Forschung in diesem Bereich beleuchtet. Im folgenden Kapitel wird der Aufbau des menschlichen Ohres geschildert, sowie die neurologischen Grundlagen des Gehörsinns.

Das nächste Kapitel erklärt ausführlich die Entdeckung und charakteristischen Eigenschaften des Residualtonhörens. Dieses Phänomen wurde als erstes von Schouten detailliert beschrieben und bezeichnet die Wahrnehmung eines physikalisch nicht vorhandenen Grundtons. Damit ist gemeint, dass im Frequenzspektrum des Klanges nur die Obertöne auftreten. Der Zuhörer nimmt den Klang jedoch so wahr, als wäre der Grundton Teil des Spektrums. Der fehlende Grundton hat somit keine Auswirkung auf die Tonhöhenwahrnehmung, da an seine Stelle ein Residualton tritt. Das Residualtonphänomen kann in Verbindung mit Lautsprechern, Telefonen oder tiefen Instrumenten wie bspw. dem Fagott beobachtet werden.

Anschließend werden die beiden grundlegenden Modelle der Tonhöhenwahrnehmung (Frequenz- vs. Zeitbereichsanalyse) vorgestellt und ihre Fähigkeit, Phänomene des menschlichen Gehörs zu erklären, bewertet.

Im nächsten Kapitel wird das Konzept der neuronalen Autokorrelation präsentiert und grundlegende mathematische Eigenschaften gezeigt. Dieser erfolgsversprechende Ansatz ist die Grundlage vieler Tonhöhenwahrnehmungsmodelle. Einige davon werden im darauffolgenden Kapitel vorgestellt und die biologischen und theoretischen Hintergründe erklärt. Forscher wie Langner konnten nachweisen, dass im auditorischen Cortex des Menschen ein Autokorrelationsprozess stattfindet. Dies spricht für ein von Licklider formuliertes und später von anderen (u.a. Ebeling, Fricke) modifiziertes Modell der Tonhöhenwahrnehmung, das sowohl auf spektrale, als auch auf zeitliche Analyse basiert und

Autokorrelation inkludiert.

Die beiden vormalig gegensätzlichen Theorien (Frequenz- vs. Zeitbereichanalyse) schaffen somit durch ihr Zusammenwirken ein neues, weitreichendes Modell zur Beschreibung der Tonhöhenwahrnehmung, dessen großer Vorteil die erstmals biologisch fundierte Erklärung des Residualtonphänomens mittels neuronaler Autokorrelation darstellt.

11.2 Abstract in English

The aim of this thesis was to describe recent and historical findings in the psychoacoustic field of pitch perception. In the beginning, an overview of historical research is given, in which the results of physicists such as Seebeck, Ohm and Helmholtz are thoroughly summarized and their importance for future research work explained. In the following chapter there is a special focus on the pitch perception phenomenon called residue. This phenomenon, first systematically described by Schouten, occurs in a situation where a listener hears a note, which only consists of the frequencies of the harmonic spectrum. That means, the fundamental frequency is not present physically, only the overtones. However, the listener is able to perceive the sound, as if the fundamental frequency were present. The residue phenomenon concerns telephone communication situations, as well as loudspeakers or instruments like the bassoon.

The two basic models of pitch perception (frequency vs. periodicity analysis) are explained, compared and their ability to account for particular pitch perception experiences is put under the microscope.

In the next chapter the idea of neural autocorrelation is presented as well as important mathematical theorems for this concept. This promising approach is the underlying idea of many current pitch perception models. Some of these models are brought into focus in the next section and again, their theoretical and biological basics are described. Recently, Langner et al. discovered biological evidence for an autocorrelation process in the auditory cortex of mammal and even human brains. This highly supports an autocorrelation model, first theoretically described by Licklider and later refined by other scientists (among others Fricke, Ebeling) as an explanation for pitch perception. The combination of the two former opposing theories (frequency vs. periodicity analysis) establishes a new, promising model for pitch perception, whose huge advantage is the fact that this model by means of neural autocorrelation even accounts for the residue phenomenon.

12 Lebenslauf der Autorin

Name Bettina Wimmer

geboren 1986 in Wels, Oberösterreich

Kontakt bettina.wimmer.mail@gmail.com

Ausbildung

Diplomstudium Musikwissenschaft an der Universität Wien absolviert (2005–2012)

Schwerpunkte: Neue Musik und Systematische Musikwissenschaft

Freie Wahlfächer aus Technischer Mathematik, Skandinavistik und Sprachwissenschaft absolviert

2008, 2009, 2010 Leistungsstipendium der Universität Wien

2008 Abschluss der 1. Diplomprüfung mit Auszeichnung

Diplomstudium Technische Mathematik an der Technischen Universität Wien (2004–laufend)

Zweig: Mathematik in Technik und Naturwissenschaften

2006 Abschluss der 1. Diplomprüfung mit Auszeichnung

2004 Matura mit Auszeichnung am BG/BRG Wels

Interessensschwerpunkte

Sprachen (Englisch, Französisch, Dänisch, Spanisch, tbc)

Musik & Kunst (Musiktheater, zeitgenössische Musik)

Mathematik & Naturwissenschaften